A6

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/12, C07K 14/705, C12Q 1/68, G01N 33/68 | A1

(11) International Publication Number: WO 99/33972

(43) International Publication Date: 8 July 1999 (08.07.99)

(21) International Application Number: PCT/CA98/01193

(22) International Filing Date: 24 December 1998 (24.12.98)

(30) Priority Data: 08/997,803 24 December 1997 (24.12.97) US

(71) Applicant (*for all designated States except US*): ALLELIX BIOPHARMACEUTICALS INC. [CA/CA]; 6850 Goreway Drive, Mississauga, Ontario L4V 1V7 (CA).

(72) Inventors; and
(75) Inventors/Applicants (*for US only*): MUNROE, Donald, G. [CA/CA]; 27 Wakefield Lane, Waterdown, Ontario L0R 2H3 (CA). GUPTA, Ashwani, K. [CA/CA]; 7031 Dunrobin Way, Mississauga, Ontario L5N 6Y4 (CA). VYAS, Tejal, B. [CA/CA]; 275 Riel Drive, Mississauga, Ontario L5B 3K1 (CA). CHUN, Jerold, J., M. [US/US]; 5747 Baja Mar, La Jolla, CA 92037 (US).

(74) Agents: CHARI, Santosh et al.; Orange & Associates, Toronto Dominion Bank Tower, Suite 3600, Toronto-Dominion Centre, P.O. Box 190, Toronto, Ontario M5K 1H6 (CA).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: MAMMALIAN EDG-5 RECEPTOR HOMOLOGS

(57) Abstract

The present invention is directed to nucleic acid sequences and amino acid sequences for mammalian EDG-5 receptor homologs, and particularly for human EDG-5 receptor homologs. The invention also provides methods for determining agonists and antagonists for EDG-5 receptors in addition to assays, expression vectors, host cells and methods for treating disorders associated with aberrant expression or activity of EDG-5. S1P and SPC are agonists for EDG-5 receptors.

MAMMALIAN EDG-5 RECEPTOR HOMOLOGS

FIELD OF THE INVENTION

The present invention is in the field of molecular biology; more particularly, the present invention describes a nucleic acid sequence and an amino acid sequence for novel mammalian, including human, EDG-5 receptor homologs.

BACKGROUND OF THE INVENTION

The family of edg receptors are commonly grouped with orphan receptors because their endogenous ligands are not known (for example see Hla T and Maciag T (1990) J Biol. Chem. 265:9308-13; US 5,585,476). Recently, however, lysophospatidic acid has been demonstrated to be the endogenous ligand for the edg-2 receptor (Hecht et al. (1996) J. Cell. Biol. 135: 1071-1083; An et al. (1997) Biochem. Biophys. Res. Comm. 213: 619-622).

The edg family of receptors is seven transmembrane G protein coupled receptors (T7Gs). T7Gs are so named because of their seven hydrophobic domains, which span the plasma membrane and form a bundle of antiparallel α helices. These transmembrane segments (TMS) are designated by roman numerals I-VII and account for structural and functional features of the receptor. In most cases, the bundle of helices forms a binding pocket; however, when the binding site must accommodate more bulky molecules, the extracellular N-terminal segment or one or more of the three extracellular loops participate in binding and in subsequent induction of conformational change in intracellular portions of the receptor. Specifically: the TM-VII is generally a highly conserved portion of the T7G receptors, and is often critically involved in ligand binding and receptor activation. the intracellular carboxy-terminal is involved in interactions with intracellular proteins, including those that transduce intracellular signals upon receptor activations; the carboxy-terminal is usually hydrophilic and highly antigenic relative to the receptor polypeptide as a whole and shows greatly reduced conservation.

Once the receptor is activated, it interacts with an intracellular G-protein complex which mediates further intracellular signaling activities generally the production of second messengers such as cyclic AMP (cAMP), phospholipase C, inositol triphosphate, activation of protein kinases, alteration in the expression of specific genes.

T7G receptors are expressed and activated during numerous developmental and disease processes. Identification of a novel T7G receptor provides the opportunity to diagnose or intervene in such processes, and the receptor can be used in screening assays to identify physiological or pharmaceutical molecules, which trigger, prolong or inhibit its activity or differentially modulate distinct intracellular pathways that are controlled from T7G receptors.

SUMMARY OF THE INVENTION

The invention provides isolated and unique nucleotide sequences which encode novel mammalian receptor homologs EDG-5, including murine EDG-5 (MEDG-5) and human EDG-5 (HEDG-5-5). Herein, the nucleotide sequence encoding MEDG-5 and HEDG-5 is designated medg-5 and hedg-5, respectively.

The present invention also relates to the isolated and unique nucleotide sequences of the complement of edg-5 mRNA, particularly hedg-5. In addition, the invention features nucleotide sequences which hybridize under stringent conditions to edg-5, particularly, hedg-5.

In addition, the present invention relates to expression vectors and host cells comprising such edg-5 nucleotide sequences.

More particularly, the present invention provides fragments which are useful as antibodies for EDG-5, for example fragments in the TM-VII and carboxy-terminal domain.

Furthermore, the invention relates to the use of nucleic acid and amino acid sequences of mammalian EDG-5, and more particularly, to the use of HEDG-5, or its variants, in the diagnosis or treatment of diseased cells and/or tissues associated with aberrant expression of HEDG-5.

Additional aspects of the invention are directed to the EDG-5 receptor, but more particularly, the HEDG-5 receptor, and include: the antisense DNA of edg-5/hedg-5; cloning or expression vectors containing edg-5/hedg-5; host cells or organisms transformed with expression vectors containing edg-5/hedg-5; chromosomal localization of hedg-5; expression

and tissue distribution of edg-5/hedg-5; a method for the production and recovery of purified EDG-5/HEDG-5 from host cells; purified protein, EDG-5/HEDG-5, which can be used to identify inhibitors for the downregulation of signal transduction involving EDG-5/HEDG-5; and methods of screening for ligands of edg-5/hedg-5 using transformed cells.

Particularly there is provided an isolated nucleotide sequence selected from the group consisting of:

(a) the nucleotide sequence comprising nucleotides 36-10974 of SEQ. ID NO: 13 (Figure 3A) (b) the nucleotide sequence of Figure 3B;

(c) the nucleotide sequence of Figure 3C;

(d) the nucleotide sequence comprising at least about 70% sequence identity to (a), (b) or (c), more preferably at least about 80-85% sequence identity, and even more preferably at least about 90% sequence identity, and most preferably at least about 95% sequence identity, and which nucleotide sequence hybridizes under stringent conditions to the nucleotide sequence of (a), (b) or (c), respectively; or portions thereof, and

(e) the nucleotide sequence which encodes the amino acid sequence of Figure 4A (SEQ ID NO. 14), 4B or 4C. There is also provided: expression vectors; host cells; purified amino acid sequences; complementary nucleic acid sequences; biologically active fragments; and hybridization probes, for such nucleotide sequences and their encoded amino acid sequences.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1A shows a partial DNA sequence of clone 501 which is a murine edg-5 clone (SEQ ID NO: 3).

Figure 1B shows the full length DNA sequence of the α subclone of a mhedg-5 pBluescript subclone and the predicted amino acid sequence thereof.

Figure 2 shows the amino acid sequence encoded by the DNA sequence of Figure 1A (SEQ ID NO: 15)

Figure 3A shows a nucleotide sequence of hedg-5 cDNA inserted into pcDNA3, nucleotides 36-1097 of which encode the full length HEDG-5. (pC3-hEdg5-3)

Figure 3B shows a nucleotide sequence of hedg-5 cDNA of clone pC3-hEdg5#3.4, which encodes the full length HEDG-5.

Figure 3C shows a nucleotide sequence of hedg-5 cDNA of clone pc3-hEdg5#28, which encodes the full length HEDG-5.

Figure 4A shows an alignment of the genomic DNA of Figure 3A (which corresponds to the cDNA of the pC3-hEdg5-3 from nt 251-1523 and the genomic DNA flanking from nt 1-250) with the predicted amino acid sequence.

Figure 4B shows the predicted amino acid sequence of hedg-5 cDNA of Figure 3B.

Figure 4C shows the predicted amino acid sequence of hedg-5 cDNA of Figure 3C.

Figure 5A shows the alignment of the predicted amino acid sequences of HEDG5 translation products of clones pC3-hedg5-3 , pC3-hedg5#4, and pC3-hedg5#28 as set out in Figures 4A, 4B and 4C, respectively.

Figure 5B shows the alignment of the amino acid sequence of murine edg-5 with the amino acid sequence of human edg-5 from the pC3-hEdg5#3.4 clone.

Figure 6 shows the functional response of the pC3-hedg5#4, pC3-hedg5-3 and pC3-hedg5#28 clones to anandamide and to LPA by activation of NF-kB production.

Figure 7 shows the SRE response and AP-1 response of pC3-hedg5#28 when treated with 10 μM LPA.

DETAILED DESCRIPTION OF THE INVENTION

The invention relates in one respect to polynucleotides, in their isolated form, that code for mammalian, including murine and human, EDG-5 receptors. The EDG receptors are characterized by structural features common to the G-protein coupled receptor class,

including seven transmembrane regions, and by the functional properties of binding lysophospholipid selectively. When expressed functionally in a host cell, i.e., in operable linkage with a responsive second messenger system the EDG-5 receptors are capable further of responding to lysophospholipid binding by signal transduction. In this regard, the activity of a G-protein coupled receptor such as an EDG-5 receptor can be measured using any of a variety of appropriate functional assays described hereinbelow.

As used herein and designated by the upper case abbreviation, EDG-5, refers to mammalian EDG-5 receptor homolog in either naturally occurring or synthetic form and active fragments thereof and the lower case edg-5 to the nucleotide sequence thereof The mammalian receptor, EDG-5, are characterised by structural features common to the G-protein coupled receptors, including the seven transmembrane regions, and by the sequence identity to each other of greater than about 56%, more preferably greater than about 70% identity, and most preferably greater than about 80% identity.

Furthermore, as used herein, the human EDG-5 receptor is designated as HEDG-5 and the nucleotide sequence as hedg-5 and the murine EDG-5 receptor is designated as MEDG-5 and the nucleotide sequence as medg-5.

The novel murine hedg-5 sequence was isolated following PCR from a murine neuronal cell line using degenerate primers based on conserved regions of transmembrane domains (TM-2) and TM-7 of the G protein-coupled receptor (GPCR) superfamily. Sequence comparison with known sequences demonstrated that this mouse clone represented a gene related to, but not identical to edg-2, an orphan GPCR. Sequence identity was 49% at the nucleotide level. In the studies detailed herein the hedg-5 sequence was used, however, these studies and the applications detailed herein could be undertaken using the novel mouse edg-5 sequence disclosed herein.

A full-length mouse sequence is obtained using methods well known to those of skill in the art. For example, by screening an arrayed mouse library (Genome Systems Inc.) using the full-length human edg-5 cDNA. The hedg-5 sequence is first radiolabelled using the condon priming method and then hybridized to the PAC filters and washed at high stringency, with the final wash done for 30 min at 65 C in 1X SSC. Genomic DNA inserts

from the clones with the strongest signals can be shotgun subcloned into pBluescript or a comparable cloning vector, using at least 3 different restriction digests of which 1 should have a 4 bp recognition site. Each digest yields a different subclone library, which in turn can be screened with the same cDNA probe under the same stringency conditions. Positives are picked, grown, mapped by restriction digest and Southern blotting to identify the size of the hybridizing insert, then sequenced using primers based on either the vector sequence, or on human edg-5 sequences. The position of the single intron seen in the human edg-5 gene should be conserved in the mouse gene. Thus, primers can be designed with a high degree of confidence to obtain the complete coding sequence of the mouse edg-5 gene without including intron sequences. Once the coding region has been determined, new PCR primers can be designed to amplify the cDNA directly from various tissue and/or cell line sources. A more detailed description of this approach can be found in Maniatis et al. Molecular Cloning, a Laboratory Manual (Cold Spring Harbor Press, 1989).

All publications and patent applications mentioned herein are incorporated by reference for the purpose of describing the methodologies, cell lines and vectors, among other things. However, nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure, for example, by virtue of prior invention.

Definitions

The following definitions are used herein for the purpose of describing particular terms used in the application. Any terms not specifically defined should be given the meaning commonly understood by one of oridinary skill in the art to which the invention pertains.

As used herein "isolated" means separated from polynucleotides that encode other proteins. In the context of polynucleotide libraries, for instance, a EDG-5 receptor-encoding polynucleotide is considered "isolated" when it has been selected, and hence removed from association with other polynucleotides within the library. Such polynucleotides may be in the form of RNA, or in the form of DNA including cDNA, genomic DNA and synthetic DNA.

As used herein "purified" refers to sequences that are removed from their natural environment, and are isolated or separated, and are at least 60% free, preferably 75 % free,

and most preferably 90% freee from other components with which they are naturally associated.

An "oligonucleotide" is a stretch of nucleotide residues which has a sufficient number of bases to be used as an oligomer, amplimer or probe in a polymerase chain reaction (PCR). Oligonucleotides are prepared from genomic or cDNA sequence and are used to amplify, reveal or confirm the presence of a similar DNA or RNA in a particular cell or tissue. Oligonucleotides or oligomers comprise portions of a DNA sequence having at least about 10 nucleotides and as many as about 35 nucleotides, preferably about 25 nucleotides.

"Probes" may be derived from naturally occurring or recombinant single - or double - stranded nucleic acids or be chemically synthesized. They are useful in detecting the presence of identical or similar sequences.

A "portion" or "fragment" of a polynucleotide or nucleic acid comprises all or any part of the nucleotide sequence having fewer nucleotides than about 6 kb, preferably fewer than about 1 kb. A portion or fragment can be used as a probe. Such probes may be labeled with reporter molecules using nick translation, Klenow fill-in reaction, PCR or other methods well known in the art. To optimize reaction conditions and to eliminate false positives, nucleic acid probes may be used in Southern, Northern or in situ hybridizations to determine whether DNA or RNA encoding HEDG-5 is present in a cell type, tissue, or organ.

"Reporter" molecules are those radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents which associate with, establish the presence of, and may allow quantification of a particular nucleotide or amino acid sequence.

"Recombinant nucleotide variants" encoding HEDG-5 may be synthesized by making use of the "redundancy" in the genetic code. Various codon substitutions, such as the silent changes which produce specific restriction sites or codon usage-specific mutations, may be introduced to optimize cloning into a plasmid or viral vector or expression in a particular prokaryotic or eukaryotic host system, respectively.

"Chimeric" molecules may be constructed by introducing all or part of the nucleotide sequence of this invention into a vector containing additional nucleic acid sequence which might be expected to change any one (or more than one) of the following HEDG-5 characteristics: cellular location, distribution, ligand-binding affinities, interchain affinities, degradation/turnover rate, signaling, etc.

"Biologically Active or Active" refers to those forms, fragments, or domains of any HEDG-5 polypeptide which retain at least some of the biological and/or antigenic activities of any naturally occurring HEDG-5.

"Naturally occurring HEDG-5" refers to a polypeptide produced by cells which have not been genetically engineered and specifically contemplates various polypeptides arising from post-translational modifications of the polypeptide including but not limited to acetylation, carboxylation, glycosylation, phosphorylation, lipidation and acylation.

"Derivative" refers to those amino acid and nucleotide sequences which have been chemically modified. Such techniques for amino acid deravitives include: ubiquitination; labeling (see above); pegylation (derivatization with polyethylene glycol); and chemical insertion or substitution of amino acids such as ornithine which do not normally occur in human proteins. A nucleotide sequence derivative would encode the amino acid which retains its essential biological characteristics of the natural molecule.

"Recombinant polypeptide variant" refers to any polypeptide which differs from naturally occurring HEDG-5 by amino acid insertions, deletions and/or substitutions, created using recombinant DNA techniques. Guidance in determining which amino acid residues may be replaced, added or deleted without abolishing activities of interest may be found by comparing the sequence of HEDG-5 with that of related polypeptides and minimizing the number of amino acid sequence changes made in highly conserved regions.

Amino acid "substitutions" are conservative in nature when they result from replacing one amino acid with another having similar structural and/or chemical properties, such as the replacement of a leucine with an isoleucine or valine, an aspartate with a glutamate, or a threonine with a serine.

"Insertions" or "deletions" are typically in the range of about 1 to 5 amino acids. The variation allowed may be experimentally determined by producing the peptide synthetically or by systematically making insertions, deletions, or substitutions of nucleotides in the hedg-5 sequence using recombinant DNA techniques.

A "signal or leader sequence" can be used, when desired, to direct the polypeptide through a membrane of a cell. Such a sequence may be naturally present on the polypeptides of the present invention or provided from heterologous sources by recombinant DNA techniques.

An "oligopeptide" is a short stretch of amino acid residues and may be expressed from an oligonucleotide. It may be functionally equivalent to and the same length as (or considerably shorter than) a "fragment", "portion", or "segment" of a polypeptide. Such sequences comprise a stretch of amino acid residues of at least about 5 amino acids and often about 17 or more amino acids, typically at least about 9 to 13 amino acids, and of sufficient length to display biological and/or antigenic activity.

"Inhibitor" is any substance which retards or prevents a biochemical, cellular or physiological reaction or response. Common inhibitors include but are not limited to antisense molecules, antibodies, and antagonists.

"Standard" is a quantitative or qualitative measurement for comparison. It is based on a statistically appropriate number of normal samples and is created to use as a basis of comparison when performing diagnostic assays, running clinical trials, or following patient treatment profiles.

"Stringent conditions" is used herein to mean conditions that allow for hybridization of substantially related nucleic acid sequences. . Such hybridization conditions are described by Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Press, 1989. Generally, stringency occurs within a range from about 5 °C below the melting temperature of the probe to about 20 °C – 25 °C below the melting temperature. As understood by ordinary skilled persons in the art, the stringency conditions may be altered in

order to identify or detect identical or related nucleotide sequences. Factors such as the length and nature (DNA, RNA, base composition) of the sequence, nature of the target (DNA, RNA, base composition, presence in solution or immobilization, etc.) and the concentration of the salts and other componenets (e.g. the presence or absence of formamide, dextran sulfate and/or polyethylene glycol) are considered and the hybridization solution may be varied to generate conditions of either low or high stringency.

"Animal" as used herein may be defined to include human, domestic (cats dogs, etc.), agricultural (cows, horses, sheep, etc.) or test species (mouse, rat, rabbit, etc.).

"Nucleotide sequences" as used herein are oligonucleotides, polynucleotides, and fragments or portions thereof, and are DNA or RNA of genomic or synthetic origin which may be single or double stranded, and represent the sense or complement or antisense strands.

"Sequence identity" is known in the art, and is a relationship between two or more polypeptide sequences or two or more polynucleotide sequences, as determined y comparing the sequences, particularly, as determined by the match between strings of such sequences. Sequence identity can be readily calculated by known methods (Computational Molecular Biology, Lesk, A.M., ed., Oxford University Press, New York, 1988; Biocomputing: Informatics and Genome Projects, Smith, D.W., ed., Academic Press, New York, 1993; Computer Analysis of Sequence Data, Part I, Griffin, A.M., and Griffin, H.G., eds., Humana Press, New Jersey, 1994; Sequence Analysis in Molecular Biology, von Heinje, G., Academic Press, 1987; and Sequence Analysis Primer, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991). While there exist a number of methods to measure identity between two sequences, the term is well known to skilled artisans (see, for example, Sequence Analysis in Molecular Biology; Sequence Analysis Primer; and Carillo, H., and Lipman, D., SIAM J. Applied Math., 48: 1073 (1988)). Methods commonly employed to determine identity between sequences include, but are not limited to those disclosed in Carillo, H., and Lipman, D., SIAM J. Applied Math., 48: 1073 (1988) or, preferably, in Needleman and Wunsch, J. Mol. Biol., 48: 443-445, 1970, wherein the parameters are as set in version 2 of DNASIS (Hitachi Software Engineering Co., San Bruno, CA). Computer programs for determining identity are publicly available. Preferred computer program methods to determine identity between two sequences include, but are not limited to, GCG program

package (Devereux, J., et al., Nucleic Acids Research 12(1): 387 (1984)), BLASTP, BLASTN, and FASTA (Atschul, S.F. et al., J. Molec. Biol. 215: 403-410 (1990)). The BLASTX program is publicly available from NCBI (blast@ncbi.nlm.nih.gov) and other sources (BLAST Manual, Altschul, S., et al., NCBI NLM NIH Bethesda, MD 20894; Altschul, S., et al., J. Mol. Bio. 215: 403-410 (1990)).Computational Molecular Biology, Lesk, A.M, ed. Unless specified otherwise in the claims, the percent identity for the purpose of interpreting the claims shall be calculated by the Needleman and Wuensch algorithm with the parameters set in version 2 of DNASIS.

The present invention provides a nucleotide sequence uniquely identifying novel mammalian, including murine (MEDG-5) and human (HEDG-5), seven transmembrane receptor (T7G) or EDG-5.

Based on the homology of HEDG-5 to human edg-2 (see table 2 below) it is likely that HEDG-5 binds a ligand of the same chemical class. Edg-2 specifically binds lysophosphatidic acid (LPA) which is a phospholipid. It was determined herein that HEDG-5 also recognizes LPA as a functional agonist.

Phospholipids have been demonstrated to be important regulators of cell activity, including mitogenisis (Xu et al. (1995) J. Cell. Physiol., 163: 441-450) and apoptosis, cell adhesion and regulation of gene expression. Specifically, for example, LPA elicits growth factor-like effects on cell proliferation (Moolenar (1996) J. Biol. Chem, 270: 12949-12952) and cell migration (Imamura et al. (1993) Biochem. Biophys. Res. Comm., 193: 497-503). It has also been suggested that LPA plays a role in wound healing and regeneration (Tigyi et al. (1992) J. Biol. Chem., 267: 21360-21367). Further, considerable circumstantial evidence indicates that phospholipids may be involved in various disease states including cancer (Imamura et al., (1993) Biochem. Biophys. Res. Comm., 193: 497-503); diseases having an inflammatory component (Fourcade et al. (1995), Cell, 80(6): 919-927, including adult respiratory distress, neurodegeneration (Jalink et al. (1993) Cell Growth Differ., 4: 247-255), rheumatoid arthritis (Natiarajan et al. (1995) J. Lipid Res., 36(9): 2005-2016), psoriasis and inflammatory bowel disease. Thus, ligands for HEDG-5, including LPA, are likely to be biologically important regulators of cell activity, and therefore aberrant expression or activity of HEDG-5 is likely to be associated with a chronic or acute disease states. Further,

modulators of HEDG-5 activity are likely to be useful in treatment or prevention of such disease states.

HEDG-5 ligands, other than LPA, are likely to be found among the phospholipid class of compounds.. Therefore, in one embodiment, preferably phospholipid molecules should be screened to identify HEDG-5 ligands. Even more preferably, lysoglycerophospholipids should be screened, such as lysophosphatidylethanolamine (LPE), lysophosphatidylserine (LPS), lysophosphatidylinositol (LPI), lysophosphatidylcholine (LPC), lyso-platelet activating factor (lyso-PAF) and phosphatidic acid. These ligands can be altered to improve metabolic stability, for example, by changing ester bond at Sn-1 to an ether on by blocking the free hydroxyl group with methoxy or acetyl ester. Additional medicinal chemistry benefits may be derived from shortening the fatty acid chain or altering the positioning of the phosphate. LPA and related phospholipids have limited solubility in aqueous solution and have a tendency to be sticky. These problems may be alleviated in a number of ways. For example, preparation of fresh stock solutions (e.g., 10 mM) by dissolving the phospholipid in calcium-free PBS and fatty-acid free BSA. Other related phospholipids can be prepared, for example, in 100% ethanol or DMS0

A diagnostic test for aberrant expression of HEDG-5 can accelerate diagnosis and proper treatment of abnormal conditions of for example, the heart, kidney, lung and testis. Specific examples of conditions in which aberrant expression of HEDG-5 may play a role include adult respiratory distress, asthma, rheumatoid arthritis, cardiac ischemia, acute pancreatitis, septic shock, psoriasis, acute cyclosporine nephrotoxicity and early diabetic glomerulopathy, as well as lung damage following exposure to cigarette smoke, asbestos or silica.

The nucleotide sequences encoding EDG-5 (or their complement) have numerous applications in techniques known to those skilled in the art of molecular biology. These techniques include use as hybridization probes, use in the construction of oligomers for PCR, use for chromosome and gene mapping, use in the recombinant production of EDG-5, and use in generation of antisense DNA or RNA, their chemical analogs and the like. Uses of nucleotides encoding EDG-5 disclosed herein are exemplary of known techniques and are not intended to limit their use in any technique known to a person of ordinary skill in the art.

Furthermore, the nucleotide sequences disclosed herein may be used in molecular biology techniques that have not yet been developed, provided the new techniques rely on properties of nucleotide sequences that are currently known, e.g., the triplet genetic code, specific base pair interactions, etc.

It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of EDG-5-encoding nucleotide sequences may be produced. Some of these will only bear minimal homology to the nucleotide sequence of the known and naturally occurring EDG-5. The invention has specifically contemplated each and every possible variation of nucleotide sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the nucleotide sequence of naturally occurring edg-5, and all such variations are to be considered as being specifically disclosed.

Although the nucleotide sequences which encode EDG-5, its derivatives or its variants are preferably capable of hybridizing to the nucleotide sequence of the naturally occurring edg-5 under stringent conditions, it may be advantageous to produce nucleotide sequences encoding EDG-5 or its derivatives possessing a substantially different codon usage. Codons can be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic expression host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding EDG-5 and/or its derivatives without altering the encoded aa sequence include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

Human genes often show considerable actual polymorphism; that is, variation in nucleotide sequence among a fraction of the entire human population. In many cases this polymorphism can result in one or more amino acid substitutions. While some of these substitutions show no demonstrable change in function of the protein, others may produce varying degrees of functional effects. In fact, many natural or artificially produced mutations can lead to expressible HEDG proteins. Each of these variants, whether naturally or artificially produced, is considered to be equivalent and specifically incorporated into the present invention.

Nucleotide sequences encoding EDG-5 may be joined to a variety of other nucleotide sequences by means of well established recombinant DNA techniques (Sambrook J et al (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor NY; or Ausubel FM et al (1989) Current Protocols in Molecular Biology, John Wiley & Sons, New York City). Useful nucleotide sequences for joining to edg-5 include an assortment of cloning vectors such as plasmids, cosmids, lambda phage derivatives, phagemids, and the like. Vectors of interest include expression vectors, replication vectors, probe generation vectors, sequencing vectors, etc. In general, vectors of interest may contain an origin of replication functional in at least one organism, convenient restriction endonuclease sensitive sites, and selectable markers for one or more host cell systems.

Another aspect of the subject invention is to provide for edg-5-specific hybridization probes capable of hybridizing with naturally occurring nucleotide sequences encoding EDG-5. Such probes may also be used for the detection of similar T7G encoding sequences and should preferably contain at least 56% nucleotide identity, more preferably at least 70% identity, to edg-5 sequence. The hybridization probes of the subject invention may be derived from the nucleotide sequence presented as SEQ. ID NO:12 or from genomic sequences including promoter, enhancers, introns or 3'-untranslated regions of the native gene. Hybridization probes may be labeled by a variety of reporter molecules using techniques well known in the art. Preferably, the hybridization probes incorporate at least 15 nucleotides, and preferably at least 25 nucleotides, of the edg-5 receptor, more particualrly of the medg-5 or the hedg-5 receptor. Suitable hybridization probes would include: consensus fragments, i.e. those regions of the mouse and human edg-5 receptor that are identical (See Figure 5B); the extracellular edg-5 binding domain, the stipulated transmembrane regions and the C-terminal portion of the receptor.

It will be recognized that many deletional or mutational analogs of nucleic acid sequences for EDG-5 will be effective hybridization probes for EDG-5 nucleic acid. Accordingly, the invention relates to nucleic acid sequences that hybridize with such EDG-5 encoding nucleic acid sequences under stringent conditions.

"Stringent conditions" refers to conditions that allow for the hybridization of substantially related nucleic acid sequences. For instance, such conditions will generally allow hybridization of sequence with at least about 70% identity, preferably with at least 80-85% sequence identity, more preferably with at least about 90% sequence identity, and even more preferably with at least about 95% sequence identity. Such hybridization conditions are described by Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Press, 1989. Hybridization conditions and probes can be adjusted in well-characterized ways to achieve selective hybridization of human-derived probes. Nucleic acid molecules that will hybridize to EDG-5 encoding nucleic acid under stringent conditions can be identified functionally, using methods outlined above, or by using for example the hybridization rules reviewed in Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Press, 1989. Without limitation, examples of the uses for hybridization probes include: histochemical uses such as identifying tissues that express EDG-5; measuring mRNA levels, for instance to identify a sample's tissue type or to identify cells that express abnormal levels of EDG-5; and detecting polymorphisms in the EDG-5. RNA hybridization procedures are described in Maniatis et al. Molecular Cloning, a Laboratory Manual (Cold Spring Harbor Press, 1989). PCR as described US Patent No's. 4,683,195; 4,800,195; and 4,965,188 provides additional uses for oligonucleotides based upon the nucleotide sequence which encodes the edg-5 sequences of the invention. Such probes used in PCR may be of recombinant origin, chemically synthesized, or a mixture of both. Oligomers may comprise discrete nucleotide sequences employed under optimized conditions for identification of edg-5 in specific tissues or diagnostic use. The same two oligomers, a nested set of oligomers, or even a degenerate pool of oligomers may be employed under less stringent conditions for identification of closely related DNA's or RNA's. Rules for designing PCR primers are now established, as reviewed by PCR Protocols, Cold Spring Harbor Press, 1991. Degenerate primers, i.e., preparations of primers that are heterogeneous at given sequence locations, can be designed to amplify nucleic acid sequences that are highly homologous to, but not identical to edg-5. Strategies are now available that allow for only one of the primers to be required to specifically hybridize with a known sequence. See, Froman et al., Proc. Natl. Acad. Sci. USA 85: 8998, 1988 and Loh et al., Science 243: 217, 1989. For example, appropriate nucleic acid primers can be ligated to the nucleic acid sought to be amplified to provide the hybridization partner for one of the primers. In this way, only one of the primers need be based on the sequence of the nucleic

acid sought to be amplified. PCR methods of amplifying nucleic acid will utilize at least two primers. One of these primers will be capable of hybridizing to a first strand of the nucleic acid to be amplified and of priming enzyme-driven nucleic acid synthesis in a first direction. The other will be capable of hybridizing the reciprocal sequence of the first strand (if the sequence to be amplified is single stranded, this sequence will initially be hypothetical, but will be synthesized in the first amplification cycle) and of priming nucleic acid synthesis from that strand in the direction opposite the first direction and towards the site of hybridization for the first primer. Conditions for conducting such amplifications, particularly under preferred stringent hybridization conditions, are well known. See, for example, PCR Protocols, Cold Spring Harbor Press, 1991.

Other means of producing specific hybridization probes for edg-5 include the cloning of nucleic acid sequences encoding EDG-5 or EDG-5 derivatives into vectors for the production of mRNA probes. Such vectors are known in the art, are commercially available and may be used to synthesize RNA probes in vitro by means of the addition of the appropriate RNA polymerase as T7 or SP6 RNA polymerase and the appropriate reporter molecules.

More particularly, a method for detection of polynucleotides that hybridize with hedg-7 is exemplified in Example 4, wherein a positive test correlates to approximately at least 70% identity, and more preferably at least 80-85% sequence identity.

It is possible to produce a DNA sequence, or portions thereof, entirely by synthetic chemistry. After synthesis, the nucleic acid sequence can be inserted into any of the many available DNA vectors and their respective host cells using techniques which are well known in the art. Moreover, synthetic chemistry may be used to introduce mutations into the nucleotide sequence. Alternately, a portion of sequence in which a mutation is desired can be synthesized and recombined with longer portion of an existing genomic or recombinant sequence.

The nucleotide sequence for hedg-5 can be used in an assay to detect inflammation or disease associated with abnormal levels of HEDG-5 expression. The cDNA can be labeled by methods known in the art, added to a fluid, cell or tissue sample from a patient, and

incubated under hybridizing conditions. After an incubation period, the sample is washed with a compatible fluid which optionally contains a reporter molecule. After the compatible fluid is rinsed off, the reporter molecule is quantitated and compared with a standard as previously defined.

The nucleotide sequence for hedg-5 has been used to construct hybridization probes for mapping the native gene. The edg-5 gene was mapped to a band p22.3 of chromosome 1 using bacterial artificial chromosomes isolated (BACs), as detailed in Example 16. Thus, the invention provides expression products from this locus that hybridize with hedg-5 (SEQ ID NO:12) under stringent conditions. In situ hybridization of chromosomal preparations and physical mapping techniques such as linkage analysis using established chromosomal markers are invaluable in extending genetic maps. Examples of genetic map data can be found in the yearly genome issue of Science (e.g. 1994, 265:1981f).

New nucleotide sequences can be assigned to chromosomal subregions by physical mapping. The mapping of new genes or nucleotide sequences provide useful landmarks for investigators searching for disease genes using positional cloning or other gene discovery techniques. Once a disease or syndrome, such as ataxia telangiectasia (AT), has been crudely localized by genetic linkage to a particular genomic region, for example, AT to 1 1q22-23 (Gatti et al (1988) Nature 336:577-580), any sequences mapping to that area may represent or reveal genes for further investigation. The nucleotide sequence of the subject invention may also be used to detect differences in gene sequence between normal and carrier or affected individuals.

Nucleotide sequences encoding edg-5 may be used to produce a purified oligo - or polypeptide using well known methods of recombinant DNA technology. Goeddel (1990, Gene Expression Technology, Methods and Enzymology, Vol. 185, Academic Press, San Diego CA) is one among many publications which teach expression of an isolated nucleotide sequence. The oligopeptide may be expressed in a variety of host cells, either prokaryotic or eukaryotic. Host cells may be from the same species from which the nucleotide sequence was derived or from a different species. Advantages of producing an oligonucleotide by recombinant DNA technology include obtaining adequate amounts of the protein for purification and the availability of simplified purification procedures.

Cells transformed with DNA encoding EDG-5 may be cultured under conditions suitable for the expression of T7Gs, their extracellular, transmembrane or intracellular domains and recovery of such peptides from cell culture. EDG-5 (or any of its domains) produced by a recombinant cell may be secreted, expressed on cellular membranes, or may be contained intracellularly, depending on the particular genetic construction used. In general, it is more convenient to prepare recombinant proteins in secreted form. Purification steps vary with the production process and the particular protein produced. Often an oligopeptide can be produced from a chimeric nucleotide sequence. This is accomplished by ligating the nucleotides from edg-5 or a desired portion of the polypeptide to a nucleic acid sequence encoding a polypeptide domain which will facilitate protein purification (Kroll DJ et al (1993) DNA Cell Biol. 12:441-53).

In addition to recombinant production, fragments of EDG-5 may be produced by direct peptide synthesis using solid-phase techniques (e.g. Stewart at al (1969) Solid-Phase Peptide Synthesis, WH Freeman Co., San Francisco QA; Merrifield J (1963) J Am Chem. Soc. 85:2149-2154). Automated synthesis may be achieved, for example, using Applied Biosystems 431A Peptide Synthesizer (Foster City, CA) in accordance with the instructions provided by the manufacturer. Additionally, a particular portion of EDG-5 may be mutated during direct synthesis and combined with other parts of the peptide using chemical methods.

EDG-5 for antibody induction does not require biological activity: however, the protein must be antigenic. Peptides used to induce specific antibodies may have an aa sequence consisting of at least five amino acids (aa), preferably at least 10 aa. They should mimic a portion of the aa sequence of the protein and may contain the entire aa sequence of a small naturally occurring molecule such as EDG-5. An antigenic portion of EDG-5 may be fused to another protein such as keyhole limpet hemocyanin, and the chimeric molecule used for antibody production.

Antibodies specific for EDG-5 may be produced by inoculation of an appropriate animal with the polypeptide or an antigenic fragment. An antibody is specific for EDG-5 if it is produced against an epitope of the polypeptide and binds to at least part of the natural or recombinant protein. Antibody production includes not only the stimulation of an immune

response by injection into animals, but also analogous processes such as the production of synthetic antibodies, the screening of recombinant immunoglobulin libraries for specific-binding molecules (e.g. Orlandi R et al (1989) PNAS 86:3833-3837, or Huse WD et al (1989) Science 256:1275-1281) or the in vitro stimulation of lymphocyte populations. Current technology (Winter G and MIstein C (1991) Nature 349:293-299) provides for a number of highly specific binding reagents based on the principles of antibody formation. These techniques may be adapted to produce molecules which specifically bind EDG-5s.

An additional embodiment of the subject invention is the use of HEDG-5 specific antibodies, inhibitors, ligands or their analogs as bioactive agents to treat inflammation or disease including, but not limited to viral, bacterial or fungal infections; allergic responses; mechanical injury associated with trauma; hereditary diseases; lymphoma or carcinoma; or other conditions which activate the genes of kidney, lung, heart, lymphoid or tissues of the nervous system.

Bioactive compositions comprising agonists, antagonists, receptors or inhibitors of HEDG-5 may be administered in a suitable therapeutic dose determined by any of several methodologies including clinical studies on mammalian species to determine maximal tolerable dose and on normal human subjects to determine safe dose. Additionally, the bioactive agent may be complexed with a variety of well established compounds or compositions which enhance stability or pharmacological properties such as half-life. It is contemplated that the therapeutic, bioactive composition may be delivered by intravenous infusion into the bloodstream or any other effective means which could be used for treating problems involving aberrant expression of the Edg-5 gene.

The examples below are provided to illustrate the subject invention. These examples are provided by way of illustration and are not included for the purpose of limiting the invention.

EXAMPLES

Example 1: PCR cloning of murine edg-5 cDNA

Poly-A+ RNA was isolated from TR and TSM murine neuronal cell lines by twice selecting on oligo-dT cellulose (Pharmacia, Cat. 27-5649-01). 10.5 μg of this RNA was reverse-transcribed with oligo-dT or random hexamers as to prime the RT reaction. RNA and primers were heated to 65°C for 5 min., then cooled to room temperature. Additional reagents were added to give the following final concentrations: 50 mM Tris-Cl, pH 8.3, 6 mM $MgCl_2$, 40 mM KCl, 1 mM DTT, 1 mM each dNTP, and 1 unit/μl of Moloney murine virus RT enzyme.

First strand cDNA was amplified in PCR reactions using degenerate primers A1 (SEQ ID NO: 1) and B1 (SEQ ID NO: 2) was conducted as follows. PCR reactions used 40 ng of first strand cDNA in 10 mM Tris-Cl, pH 8.3, 50 mM KCl, 2 μM of each primer, 1.5 mM $MgCl_2$, 0.2 μM each dNTP and 2.5 units of Taq DNA polymerase. Thirty pairwise combinations of primers were used. Reactions were placed in a Perkin-Elmer 480 thermal cycler, denatured for 3 min. at 94°C, and then cycled 25-40 times at 96°C for 45 sec, 47°C for 144 sec or 53°C for 216 sec, and 72°C for 3 min. initially, increasing 6 sec/cycle. Products were cloned using the TA PCR cloning vector (Invitrogen, Cat. K2000-40). The resulting edg-5 clone, 501(SEQ ID NO: 3) was sequenced by the dideoxy termination method.

A1: 5'-AAYTRSATIMTISTIAAYYTIGCIGTIGCIGA-3' (SEQ ID NO: 1)
B1: 5'-CTGIYKWTTCATIAWIMMRTAIAYIAYIGGRTT-3' (SEQ ID NO: 2)

The nucleotide sequence of clone 501(SEQ ID NO: 3) is shown in Figure 1A. A search of Genbank showed that clone 501 (SEQ ID NO: 3) was most closely related to the LPA receptor, also identified in Genbank as the GPCR orphan edg-2 (Genbank MMU70622). Sequence identity between the clone 501 (SEQ ID NO: 3) and edg-2 was 60.5% over the 639 bp length of clone 501 (SEQ ID NO: 3). The amino acid sequence of this nucleotide is shown in Figure 2 (SEQ ID. NO: 15)

Approximately $5x10^5$ phage from an embryonic day 15 whole embryo lambda-ZAP cDNA library (Clontech) was screened with part of this PCR product, 501AB (the original PCR probe from the degenerated PCR screen, using PCR primers A and B), on conventional nylon filter lifts with ^{32}P-labeled probe, and washed to high stringency. Two clones were

isolated and subcloned into the EcoRI site of pBluescript. One of the clones was sequenced as shown in Figure 1B with the amino acid sequence shown in Figure 5B.

EXAMPLE 2: Isolation of hedg-5 cDNA PCR amplification of partial hedg-5 gene from human genomic DNA

PCR primers JC501-F2 (SEQ ID NO: 4) and JC501-R (SEQ ID NO: 5) were designed using the sequence of clone 501 (SEQ ID NO: 3) and used to obtain a PCR fragment of hedg-5, as detailed below with the Expand™ PCR system from Boehringer Mannheim (Cat. 1681-842). Human genomic DNA was obtained from Promega (Cat. G304A).

JC501-F2:
5'-TTTTTACTCGAGATTTGCTGGTTATTGCTGTGGAAAG-3' (SEQ ID NO: 4)

JC501-R:
5'-TTTTTTCTAGACGGTCATCACTGTCTTCATTAGCTTC-3' (SEQ ID NO: 5)

Each reaction contained the following reagents:

30.25 μl water
10 μl 2.5 mM dNTP mix
5 μl 10x Expand™ Buffer 3
1.5 μl 10 μM JC501-F2 primer
1.5 μl 10 μM JC501-R primer
0.75 μl Expand PCR enzyme (3.5 units/μl)
1 μl human genomic DNA (0.272 μg/μl)

PCR Conditions:
Incubate: 94°C for 2 min.
30 cycles: 92°C for 1 min.
45°C for 5 min.
68°C for 1 min.

Hold: 4°C

On ethidium bromide (EtBr)-stained agarose gel, an intense PCR product of about 390 bp was seen. This product was reamplified in the following PCR reaction:

30.25 μl water
10 μl 2.5 mM dNTP mix
5 μl 10x Expand™ Buffer 3
1.5 μl 10 μM JC501-F2 primer
1.5 μl 10 μM JC501-R primer
0.75 μl Expand PCR enzyme (3.5 units/μl)
1 μl PCR product from the previous PCR reaction

PCR Conditions:
Incubate: 94°C for 2 min.
30 cycles: 92°C for 1 min.
45°C for 5 min.
68°C for 1 min.
Hold: 4°C

The intense 390 bp product of the PCR reamplification was excised from the agarose gel. The PCR products from 30 μl of the PCR reaction were purified from pooled gel slices using a Qiaquick Gel extraction kit (Qiagen Inc.; Cat. 28706) and eluted with 20 μl of 10 mM Tris-Cl, pH 8.5. The eluted DNA was quantitated and the sequence of the PCR product was determined by automated sequencing at Allelix's in-house facility, with an ABI 377 Sequencer and fluorescent dideoxy terminators, using each primer from the PCR reactions shown above.

Sequencing results showed 81.5% identity at the nucleotide level with the sequence of mouse clone 501, over a 312 bp overlap excluding the primer sequences.

PCR amplification and sequencing of large edg-5 cDNA fragments

Primers, H501-20F (SEQ ID NO: 6) and H501-246R(SEQ ID NO: 7), specific to hedg-5, were used to amplify cDNAs encoding larger portions of hedg-5 form a λgt10 fetal heart cDNA library, as follows.

H501-20F:
5'-ATGCGGCTGCATAGCAACCTGACCAAAAAG-3' (SEQ ID NO: 6)

H501-246R:
5'-ATCCGCAGGTACACCACAACCATGATGAGG-3' (SEQ ID NO: 7)

Each reaction contained the following reagents:

30.25 μl water
10 μl 2.5 mM dNTP mix
5 μl 10x Expand™ Buffer 3
1.5 μl 10 μM H501-20F primer
1.5 μl 10 μM H501-246R primer
0.75 μl Expand PCR enzyme (3.5 units/μl)
1 μl fetal heart cDNA library (≥1 library equivalent/μl; Clontech; Cat. HL5017a)

PCR Conditions:
Incubate: 94°C for 2 min.
30 cycles: 92°C for 1 min.
45°C for 5 min.
68°C for 1.5 min.
Incubate: 68°C for 8 min.
Hold: 4°C

On EtBr-stained agarose gel, a moderately intense 250 bp PCR product was seen in a fetal heart library, the approximate size expected from the positions of the primers. No specific PCR products were seen in any of 13 other cDNA libraries tested.

To obtain additional edg-5 sequence, and possibly amplify the full-length cDNA from the fetal heart cDNA library, PCR reactions were conducted using JC501-F2 (SEQ ID NO: 4) or JC501-R (SEQ ID NO: 5) primers versus primers derived from the lgt10 vector in which this cDNA library was constructed. Although cDNA inserts are not directionally cloned into the lgt10 vector, we chose to amplify products only from one direction. The vector-based primer sequences were:

GT10-F: 5'-TTTTGAGCAAGTTCAGCCTGGTTAAGT-3' (SEQ ID NO: 8)
GT10-R: 5'-TGGCTTATGAGTATTTCTTCCAGGGTA-3' (SEQ ID NO: 9)

One PCR reaction was done with JC501-F2 (SEQ ID NO: 4) vs. GT10-R (SEQ ID NO: 9) primers to amplify the 3' end of edg-5 cDNA clones, and another was done with GT10-F (SEQ ID NO: 8) vs. JC501-R (SEQ ID NO: 5) primers to amplify the 5' end of edg-5 cDNA clones. Each 40 μl reaction contained the following reagents:

23.6 μl water
8.0μl 2.5 mM dNTP mix
4 μl 10x Expand™ Buffer 3
2.0 μl 10 μM edg-5 specific primer
0.8 μl 10 μM vector primer
0.6 μl Expand PCR enzyme (0.4 units)
1 μl cDNA library stock (≥1 library equivalent/μl; Clontech; Cat. HL5017a)

PCR Conditions:
Incubate: 94°C for 2 min.
30 cycles: 92°C for 30 sec
55°C for 2 min.
68°C for 3 min.
Incubate: 68°C for 8 min.
Hold: 4°C

The results showed 2 faint PCR products (designated 510-5-1 and 510-5-2) from the 3'-end PCR reaction (JC501-F2 (SEQ ID NO:4 /GT10-R (SEQ ID NO:9). From the 5'-end PCR reaction (GT10-F (SEQ ID NO: 8)/JC501-R (SEQ ID NO:5) again 2 faint PCR bands (designated 510-6-1 and 510-6-2) were seen. Each band was tip-eluted from the gel by stabbing the band with a fresh Pipetman plugged tip, which was then rinsed into 50 µl of TE, pH 8. This solution was used as a stock from which nested reamplifications were done, using the same vector primer vs. a nested human-specific primer as follows:

11.5 µl water
4.0µl 2.5 mM dNTP mix
2 µl 10x Expand™ Buffer 1
0.6 µl 10 µM edg-5 specific primer
0.6 µl 10 µM vector primer
0.3 µl Expand PCR enzyme (0.4 units)
1 µl tip-eluted PCR DNA stock

PCR Conditions:
Incubate: 94°C for 2 min.
30 cycles: 92°C for 30 sec
55°C for 40 sec
68°C for 3 min.
Incubate: 68°C for 8 min.
Hold: 4°C

DNA from the most intense band of each nested reamplification was purified using a QIAquick Gel extraction kit and eluted in 50 ul of 10 mM Tris-Cl, pH 8.5.

Full-length cloning of the hedg-5 cDNA into pcDNA3 vector

Extension PCR (cycles without primers) was used to extend the overlapping ~1.0 kb 3' fragment (designated 511-5: reamplified from 510-5-2) and 700 bp 5' fragment (designated 511-14: reamplified from 510-6-2) as follows:

Extension PCR :
19.8µl water
5.6 µl 2.5 mM dNTP mix
4.0 µl 10x Expand™ Buffer 1
5 µl edg-5 3' PCR DNA fragment (511-5)
5 µl edg-5 5' PCR DNA fragment (511-14)
0.6 µl Expand PCR enzyme (3.5 units/µl)

PCR Conditions:
Incubate 94°C for 2 min.
15 cycles: 92°C for 1 min.
60°C for 10 min.
68°C for 3.5 min.
Incubate: 68°C for 8 min.
Hold: 4°C

Two microliters of the extension PCR reaction was then reamplified using the two vector primers (GT10-F (SEQ ID NO:8) and GT10-R (SEQ ID NO:9) to select for full-length extension products.

32.25 µl water
7.0 µl 2.5 mM dNTP mix
5.0 µl 10x Expand™ Buffer 1
1.5 µl 10 µM GT10-F primer
1.5 µl 10 µM GT10-R primer
0.75 µl Expand PCR enzyme (3.5 units/µl)

PCR Conditions:
Incubate 94°C for 2 min.
30 cycles: 92°C for 40 sec
50°C for 40 sec
68°C for 3 min.
Incubate: 68°C for 8 min.

Hold: 4°C

After gel electrophoresis of the PCR products, an intense DNA band of about 1.4 kb was seen. The PCR product was purified with a QIAquick PCR purification kit (QIAGEN Inc., Cat. 28106), eluted in 50 µl of 10 mM Tris-Cl, pH 8.5. The gel-purified PCR fragment was then sent for automated sequencing at Allelix's in-house facility, as described above. The sequencing results confirmed the identity of the amplified band as edg-5, and suggested that a full-length clone of edg-5 had been reconstructed by extension PCR.

To subclone into pcDNA3 the above DNA was re-amplified with modified vector primers GT10-5KXb (SEQ ID NO: 10) and GT10-3BXh (SEQ ID NO: 11).

GT10-5KXb :
5'-GGGTAGTCGGTACCTCTAGAGCAAGTTCAGCC- 3' (SEQ ID NO: 10)
GT10-3BXh :
5'- ATAACAGAGGATCCTCGAGTATTTCTTCCAG- 3' (SEQ ID NO: 11)

Reamplification PCR:

67.5 µl water
14 µl 2.5 mM dNTP mix
10 µl 10x Expand™ Buffer 1
3 µl 10 µM GT10-5KXb primer
3 µl 10 µM GT10-3BXh primer
1.5 µl Expand PCR enzyme (3.5 units/µl)
1 µl DNA from previous PCR reaction

PCR Conditions:

Incubate 94°C for 2 min.
5 cycles: 92°C for 1 min.
50°C for 1 min.
68°C for 2 min.
25 cycles: 92°C for 1 min.
60°C for 1 min.

68°C for 2 min.
Incubate: 68°C for 8 min.
Hold: 4°C

The PCR product was QIAquick PCR-purified and eluted in 50 µl of 10 mM Tris-Cl, pH 8.5 as described previously and restricted with KpnI and XhoI.

Restriction digest of PCR sample with KpnI and XhoI:
Two successive restriction digests was performed on the purified extension PCR product as follows:

38 µl Extension PCR DNA
5 µl 10X NEBuffer 1 (New England Biolabs [NEB])
2 µl KpnI restriction endonuclease (10 units; NEB, Cat #142S)
5 µl 10X Acetylated BSA stock (NEB)

The restriction digest was incubated for 1 hour in a 37°C water bath. and then the following reagents and enzyme were added:

10 µl 10X NEBuffer 2 (NEB)
1 µl XhoI restriction endonuclease (20 units; NEB, Cat #146S)
5 µl 10X Acetylated BSA stock (NEB)
43 µl water

The reaction products were purified using a QIAquick PCR purification kit and eluted in 50 µl of 10 mM Tris-Cl, pH 8.5.

Preparation of pcDNA3 cloning vector with KpnI and XhoI:

4 µl pcDNA3 plasmid DNA (Invitrogen; Cat. V790-20) containing a 1.8 kb cDNA insert
10 µl 10X NEBuffer 2 (NEB)
3 µl KpnI restriction endonuclease (NEB: 1:10 dilution; 3 units)

3 µl XhoI restriction endonuclease (NEB: 1:20 dilution; 3 units)
10 µl 10X Acetylated BSA stock (NEB)
64 µl water

The vector DNA was digested for 1 hour at 37°C. Then, 3 units more of each enzyme was added and the tubes were incubated for a further 2 hr. The digest was run on a gel, and the vector DNA band without cDNA insert was excised, purified using GeneClean II kit (BIO 101) and eluted in 40 µl of 10 mM Tris-Cl, pH 8.5.

The double-digested, gel-purified PCR DNA was ligated into the prepared pcDNA3 plasmid vector using T4 DNA ligase kit (NEB, Cat. 202CS) and transformed into Epicurean Coli XL-2 Blue MRF' Ultracompetent cells (Stratagene, Cat. 200150). The transformation was plated onto 2xYT/Ampicillin plates and single colonies were picked. DNA minipreps were made using QIAGEN QIA-Prep8 miniprep kit (Cat. 27144) and clones with appropriate inserts were identified by sequencing, carried out with the in-house ABI automated sequencing system. From this analysis, a clone designated pC3-hedg-5-3 (SEQ ID NO:13) was chosen for complete sequence determination of the cDNA insert.

Features of the hedg-5 cDNA

A BLAST search of Genbank, EMBL, dbEST, and the GSS and STS genomic sequencing databases indicates that the hedg-5 sequence is novel. The bovine LPA receptor, edg-2, was the highest-scoring full-length cDNA sequence found from the combined Genbank/EMBL databases (Genbank BTU48236: 55% identity).

This sequence includes 10 bp of 5'-untranslated sequence, the edg-5 open reading frame of 1059 bp, and a 3'-untranslated region spanning 204 bp. The coding region of edg-5 begins with the first methionine codon, at nt 36-38 and terminates with the stop codon at nt 1095-1097. The prediction of this open reading frame is supported by the sequence of genomic DNA flanking the 5' end of the cDNA sequence (see below). 250 bp of 5' flanking sequence was obtained from a BAC genomic clone as described in Example 16 (Figure 4A, SEQ ID NO: 12). The proposed translation start site was preceded by an in-frame stop codon 24 bp upstream. Sequencing of different clones revealed the existence of several

sequence polymorphisms, which may represent a sampling of natural variability of the edg-5 sequence within the human population. The 15 polymorphisms observed within the edg-5 open reading frame are listed below. Nine of these substitutions did not result in a change in the encoded amino acid), while 3 resulted in conservative substitutions and 3 resulted in nonconservative substitutions.

Table 1. Apparent polymorphisms in the hedg-5 protein coding region.

Nucleotide Position	Affected Codon & Polymorphism	Amino Acid Predicted	Consequence
491	TTC	Phenylalanine	
	TTT	Phenylalanine	Silent
585	CTG	Leucine	
	TTG	Leucine	Silent
716	GTC	Valine	
	GTT	Valine	Silent
779	ATC	Isoleucine	
	ATT	Isoleucine	Silent
781	TCT	Serine	Nonconservative
	TTT	Phenylalanine	Substitution
788	TGC	Cysteine	
	TGT	Cysteine	Silent
790	TCT	Serine	Nonconservative
	TTT	Phenylalanine	Substitution
830	TTC	Phenylalanine	
	TTT	Phenylalanine	Silent
874	GTG	Valine	Conservative
	GCG	Alanine	Substitution
887	ATC	Isoleucine	
	ATT	Isoleucine	Silent
914	AAC	Asparagine	Conservative
	AAA	Lysine	Substitution
917	GTC	Valine	
	GTT	Valine	Silent

922	TCT	Serine	Nonconservative
	TTT	Phenylalanine	Substitution
1041	CTC	Leucine	Conservative
	TTC	Phenylalanine	Substitution
1277	GAG	Glutamate	
	GAA	Glutamate	Silent

The edg-5 open reading frame of the pC3-hEdg5-3 (SEQ ID NO:13; Figure 3A) clone predicts a 353 amino acid polypeptide (SEQ ID NO: 14, Figure 4A) with many typical features of a GPCR. These include:

1. A hydropathy profile consistent with the 7 transmembrane structure of GPCRs:
   - N-terminal extracellular domain: 1-30
   - TM-1: 31-56
   - IL-1: 57-63
   - TM-2: 64-92
   - EL-1: 93-106
   - TM-3: 107-125
   - IL-2: 126-144
   - TM-4: 145-170
   - EL-2: 171-186
   - TM-5: 187-207
   - IL-3: 208-239
   - TM-6: 240-261
   - EL-3: 262-276
   - TM-7: 277-297
   - C-terminal cytoplasmic domain: 298-353
2. Potential N-glycosylation site in the extracellular N-terminal domain, at residue 15
3. Potential protein kinase C phosphorylation sites at residues 141, 229 and 303
4. Potential cAMP- and cGMP-dependent kinase phosphorylation sites at residues 217, 233 and 321
5. Potential casein kinase-II phosphorylation site at residue 329

The amino acid sequence of the human edg-5 receptor, SEQ ID NO:14 (Figure 4A), also shows high homology with other members of the edg subfamily of GPCRs. . The pairwise percent identity and similarity is presented in Table 2 below:

Table 2.

Percent Amino Acid Identity and Similarity of Edg Family Sequences to the Human Edg-5 receptor

Gene	Percent Identity	Percent Similarity
Edg-1 (Human)	30.1	40.9
Edg-2 (Human)	48.6	59.0
Edg-2 (Bovine)	55.1	
Edg-3 (Human)	32.6	43.3
H218 (Edg-4 - Rat)	31.6	40.6
Edg-6 (Human)	46.0	55.5

Multiple sequence alignment indicates that edg-2 is the closest known relative of edg-5 at the amino acid sequence level, as suggested by the DNA sequence. The edg-5 gene product is also closely related to edg-6, a novel edg gene described in copending application USSN 08/763,938. Edg-2, edg-5 and edg-6 appear to form a subfamily distinct from edg-1, edg-3 and edg-4 within the larger edg gene family.

Alternative splicing variants of murine edg-2 have been found, which differ in length within the N-terminal coding region. The longer open reading frame (Genbank, accession no:MMU70622) encodes an 18-amino acid N-terminal extension of the shorter open reading frame (Genbank, accession no:MMU48235), and retains the initiator methionine codon of the shorter product as amino acid 19 of the longer product. Due to the sequence relatedness of edg-2 and edg-5 and the fact that the methionine codon of the shorter edg-2 product aligns closely with the initiation codon of hedg-5, the edg-5 open reading frame hedg-5 may encode a similar N-terminal extension to the HEDG-5 peptide of SEQ ID NO:14. Such an extension will result from splicing of sequences found upstream of the hedg-5 sequences presented herein, and will produce one or more spliced mRNA variants with a N-terminal extensions. Briefly, given the instant disclosure the skilled artisan could discover such splice variants by 5' RACE using a commercially available 5' RACE kit (Life Technologies, Cat No:18374-041) using the approach detailed in start protocols in Molecular Biology (2nd edition, 15-27). Briefly, first strand cDNA is primed using an antisense oligonucleotide specific for hedg-5

and ideally directed to a sequence about 500 nucleotides from the 5' end of the known hedg-5 sequence; kidney and lung RNA are preferred templates for cDNA synthesis. Thereafter, first strand cDNA is then tailed using terminal transferase, for example, with deoxyguanine residues. PCR amplification is primed using an anchor primer complementary to the polyguanine tail and a nested primer specific to hedg-5.

EXAMPLE 3: Molecular cloning of hedg5 coding region for expression and functional analysis in eukaryotic cells.

After surveying various cDNA libraries and first strand cDNA preparations, we were unable to obtain a full-length clone. The rarity of edg-5 in cDNA libraries is further supported by a complete absence of EST's from the edg-5 coding regions in the DBEST database, which contains millionsof individual EST's. Therefore, an alternative approach was designed. In this approach, the coding region would be amplified in two fragments from genomic DNA, since we previously determined the location of the single splice site that occurs (between nt 771/772 of SEQUENCE ID NO: 13) in the genomic DNA encoding HEDG5. Then, the two fragments would be joined by an extension PCR in which primers were engineered to contain a 30 bp overlap.between the two fragments to obtain a functional, full-length edg5 cDNA, DNA fragments from two exons next to intron located at nt 996/997 were PCR amplified using the following primers so that they have an overlap of 30 nt.

5' Exon Fragment

HE5-261F: [5'-ATGAATGAGTGTCACTATGACAAG-3']
HE5-1011R: [5'-ATACCACAAACGCCCCTAAGACAGTCATCACCGTCTTC-3']

3' Exon Fragment

HE5-982F: [5'-TGATGACTGTCTTAGGGGCGTTTGTGGTATGCTGGACC-3']
HE5-1322R: [5'-TTAGGAAGTGCTTTTATTGCAGACTGC-3']

Human genomic DNA (Clontech, Cat #6550-1) was amplified with each pair of primers under the following condition of PCR amplification by using Expand™ PCR system from Boehringer-Mannheim (Cat.1681-842).

Each reaction contained the following reagents:

5.0 µl 10x PCR Buffer 3
1.0 µl 25 mM dNTP mix
1.5 µl Primer HE5-261F or HE5-982F (10 pmol/l)
1.5 µl Primer HE5-1011R or HE5-1322R (10 pmol/l)
0.75 µl Expand™ Enzyme (7.5 units)
38.25 µl water
2.0 µl DNA

PCR conditions:

Incubate:	94°C for 2 min
30 cycles:	94°C for 1 min
	55°C for 2 min
	68°C for 1 min
Incubate:	68°C for 8 min
Hold:	4°C

DNA fragments of approximately 700 bp (5' exon) and 350 bp (3' exon) were amplified. The two DNA fragments were purified using Qiaquick gel extraction kit (Qiagen, Cat. 28706) and eluted in 50 ul of 10 mM Tris, pH 8.5. Extension PCR (cycles without primers) was then used to join the 5' exon and 3' exon fragments, which overlapped each other by 30 bp.

Extension PCR:

Each reaction contained the following reagents:

2.0 µl 10x PCR Buffer 3

0.4 µl 25 mM dNTP mix
0.3 µl Expand™ Enzyme (2.5 units)
13.8 µl water
1.0 µl 5' exon PCR-amplified DNA
1.0 µl 3' exon PCR-amplified DNA

PCR conditions:

Incubate:	94°C for 2 min
20 cycles:	94°C for 1 min
	60°C for 5 min
	68°C for 1.5 min
Incubate:	68°C for 8 min
Hold:	4°C

Five µl of the amplified product from the above PCR was then reamplified under the following condition of PCR with primers HE5-261F and HE5-1322R, described previously.

Each reaction contained the following reagents:

5.0 µl 10x PCR Buffer 3
1.0 µl 25 mM dNTP mix
1.5 µl Primer HE5-261F (10 pmol/µl)
1.5 µl Primer HE5-1322R (10 pmol/µl)
0.75 µl Expand™ Enzyme (7.5 units)
35.25 µl water
5.0 µl DNA

PCR conditions:

Incubate:	94°C for 2 min
25 cycles:	94°C for 1 min
	55°C for 1 min
	68°C for 1 min
Incubate:	68°C for 8 min

Hold: 4°C

An intense DNA band of about 1.0 kb was purified using the Qiaquick PCR purification kit (Qiagen), eluted in 50 µl of 10 mM Tris, pH 8.5 and was sent for direct PCR sequencing with each primer used in the above PCR reactions, as described previously. The resulting sequences showed 93 – 99% identity to human edg5 cDNA, within the edg-5 coding region.

To subclone into pcDNA3.1 (Invitrogen; Cat. V795-20) the above DNA was reamplified with modified primers HE5-KZKF and HE5-Kpn1322R under the following conditions:

HE5-KZKF: [5'-TTTAAACTCGAGCCACCATGAATGAGTGTCACTATGAC - 3']
HE5-Kpn1322R: [5'-TATATAGGTACCTTAGGAAGTGCTTTTATTGCAGACTGC-3']

Each reaction contained the following reagents:

5.0 µl 10x PCR Buffer 3
1.0 µl 25 mM dNTP mix
1.5 µl Primer HE5-KZKF (10 pmol/µl)
1.5 µl Primer HE5-Kpn1322R (10 pmol/µl)
0.75 µl Expand™ Enzyme (7.5 units)
39.25 µl water
1.0 µl DNA

PCR conditions:

Incubate:	94°C for 2 min
25 cycles:	94°C for 40 sec
	50°C for 1 min
	68°C for 1.5 min
25 cycles:	94°C for 40 sec

	65°C for 40 sec
	68°C for 1.5 min
Incubate:	68°C for 8 min
Hold:	4°C

The PCR product was purified as described previously and subcloned into Kpn1 and Xho1 restriction sites of pcDNA3.1.

Plasmid DNA was prepared from several positive clones and cotransfected into 293-EBNA cells together with the 2xSRE-Luciferase reporter plasmid.

Transient transfection protocol for 293-EBNA:

Day 1.

1) 100 mm plates of 293-EBNA with a confluency of ~80% were used for transfection.

2) NF-kB Reporter Gene Cotransfection: Expression plasmid (3.5 μg) and reporter plasmid (6XNF-kBtk-p4Luc-zeo; 0.5 μg) DNA samples were combined and diluted in 750 μl of DMEM/F12 (serum-free media) and 20 μl Plus Reagent (Lipofectamine Plus Kit, Life Technologies Cat. 10964-013), and incubated at room temperature for 15 min.

3) 30 μ Lipofectamine Reagent (Lipofectamine Plus Kit) was diluted in 750 μl DMEM/F12. The diluted Lipofectamine was then combined with the DNA/Plus mixture and incubated at RT for 15 min.

4) The 293-EBNA plates were washed once with PBS, and 5 ml DMEM/F12 was added to each plate.

5) DNA/Plus/Lipofectamine mixture was added to each plate of 293-EBNA cells. The plates were left for 3 hr at 37°C in a 5% CO_2 incubator.

6) The transfection medium was replaced with DMEM/F12 containing 10% FBS to recover overnight.

Day 2.

1) Transfected cells were harvested by trypsinization and 20,000 cells per well were plated in 96-well Blackview plates coated with poly D-lysine (Becton Dickinson Labware, Cat.

40640). Medium was DMEM/F12 containing 0.5% FBS. No cells were plated in the outside wells of the 96-well plate. Cells were returned to the incubator for 24 hr.

Day 3.

1) Media was removed and cells treated with compounds diluted in DMEM/F12 media containing 0.15% FBS and the following treatments: a) Untreated: DMEM/F12 plus 0.15% FBS; b) AN (10 μM anandamide); c) LPA (10 μM oleyl lysophosphatidic acid).

2) The cells were treated overnight in the incubator.

Day 4.

1) Luclite kit (Packard; Cat. 6016911) was used for luciferase assay. All reagents were brought to room temperature before use.
2) Media was removed from each well. 50 μl 0.5M HEPES pH 7.8, 1 mM $MgCl_2$, 1 mM $CaCl_2$ was added to all wells of 96-well plate.
3) Luclite substrate was made up and 50 μl substrate was added to each well as specified in the kit.
4) Plates were incubated at room temperature for 30 min.
5) After incubation, plates were counted in a 12-detector Packard Top Count on a program without dark delay.

Results:

As we have documented elsewhere (See U.S. provisional patent application entitled "Identification of Lysolipid Receptors Involved in Inflammatory Response" filed on November 25, 1998 by MUNROE and corresponding PCT application filed on December 30, 1998), edg-2, edg-5 and edg-6 proved to be inflammatory LPA receptor subtypes of the edg receptor family which when activated induce NF-kB. As exemplified in Figure 6, it was determined that HEDG-5, as particularly represented by the two clones pc3-hedg5#3-4 and pc3-hedg5#28, responded to LPA but not anandamide at 10 μM to activate NF-kB. (See Figure 6)

Three inflammatory subtypes of lysophosphatidic acid (LPA) receptor.

An additional experiment was carried out to test the response of clone #28 in reporter gene constructs with the serum response element (SRE) or the proximal 1 kb of the human collagenase gene promoter containing an inducible enhancer element for activator protein-1 (AP-1) together with the edg-2 and edg-6 receptor sub-types. As shown in Figure 7, the pC3-hedg5#28 showed an SRE response and an AP-1 response when treated with 10 μM LPA.

To determine whether these receptors might mediate inflammatory responses, each was cotransfected separately with SRE, NF-kB or AP-1 reporter genes. The AP-1 reporter contained approximately 1 kb of the human collagenase II promoter, and the first 50 bp of the 5'-untranslated region of the collagenase II transcription unit (Angel P, et al. Phorbol ester-inducible genes contain a common cis element recognized by a TPA-modulated trans-acting factor. Cell. 1987 Jun 19;49(6):729-739.), a region whose inducible expression has been shown to be controlled by AP-1. This transcription factor, like NF-kB has been implicated in inflammatory and neoplastic signal transduction., though the gene targets of its action are largely distinct from those of NF-kB (Adcock IM. Transcription factors as activators of gene transcription: AP-1 and NF-B. Monaldi Arch Chest Dis. 1997 Apr;52(2):178-86. Review.).

293-EBNA cells were grown, lipofected in monolayer cultures, and pretreated as described above for Example 11, assay #1, except that NF-kB and AP-1 reporter-transfected cells were pretreated for 6 hr in medium containing 0.5% FBS, then treated overnight in the same medium with or without 10μM LPA.

Results: As shown in Fig. 7, all three receptors robustly activated the NF-kB reporter (about 3-4-fold) in the presence of 10 μM LPA, while no response to LPA was seen when the NF-kB reporter was cotransfected with the empty expression vector pcDNA3. With the SRE and AP-1 reporter genes, some endogenous response to LPA was seen (about 1.5-fold vs untreated control cells). However, edg-6 strongly induced both reporters, while edg-2 and edg-5 caused greater than 2-fold induction of the SRE and AP-1 reporters with LPA. Therefore, all three LPA receptors tested here are capable of inducing inflammatory gene transcription through NF-kB, and perhaps, AP-1 as well.

EXAMPLE 4: Detection of hedg-5 polynucleotides by hybridization with hedg-5.

Edg polynucleotides can vary through the introduction of natural or artificial mutations or through cloning and subsequent manipulations. Moreover, the mammalian homolog of a given gene usually varies by 10-30% from species to species, as a result of nucleotide changes that have accumulated through their divergent evolutionary history. Therefore, a method is provided herein for the detection and identification of hedg variants and other highly related genes.

The HEDG-5 coding region of hedg-5 is prepared by restriction of either pC3-hEdg5-3 or pC3-hedg5#3.4 or pC3-hedg5#28 with appropriate restriction enzymes to release the full-length hedg-5 insert, followed by cDNA insert purification using standard techniques after agarose gel electrophoresis. The cDNA insert may be labeled using ^{32}P-nucleotide end-labeling or random priming (several kits are commercially available), or through incorporation of non-natural nucleotides for later detection with antibodies by methods well known in the art. Nylon filters (e.g. Hybond N+, Cat. RPN132B) bearing a polynucleotide or mixture of polynucleotides are prepared by standard techniques. Examples include Southern blots, filter lifts from bacterial colonies or bacteriophage plaques and the like.

The dried filters are rehydrated in water, then prehybridized in a sealable bag with 10 ml (or enough to cover filters and seal the bag) of hybridization solution (48% deionized formamide, 4.8× SSC [20× SSC is 3 M NaCl, 0.3 M sodium citrate, pH 7.0], 1× Denhardt's solution [50× Denhardt's is 1% Ficoll 400, 1% polyvinylpyrrolidone, 1% BSA (Pentax Fraction V)], 10% dextran sulfate, 0.1% sodium dodecyl sulfate [SDS]) for 1 hr or more at 42°C.

Radiolabeled probe is added to 1 ml of sonicated herring sperm DNA (2 mg/ml) in a screw-cap tube and incubated in a boiling water bath for 10 min. Transfer the tube to ice, add 2 ml of hybridization solution and inject the probe solution into the sealed bag. Sufficient probe should be added to give 1 to 15 ng of radiolabeled probe/ml hybridization buffer (final volume) at $>5\times10^7$ cpm/μg DNA. Reseal the bag, mix thoroughly and incubate overnight at 42°C in a shaking or rotating water bath or incubator.

Wash filters three times with 500 ml of low-stringency wash buffer (2× SSC, 0.1% SDS) at RT for 15 min per wash, on a slowly rotating platform. Then wash two times with medium-stringency wash buffer (1× SSC, 0.1% SDS) at 65°C 15 min per wash. Dry the filters and expose to Phosphorimager cassette or autoradiography film. Positive spots or DNA bands are identified after subtraction of background or appropriate negative control samples (see below).

If needed, a DNA spot containing 10 pmol of the full-length hedg insert of pC3-hEdg5-3 can be used as a positive control (Pos) on the filter, and a DNA spot containing 10 pmol of full-length human edg-2 insert (edg-2 open reading frame only) can be used as a negative control (Neg). A full-length open reading frame, or a partial-length open reading frame, of a test DNA (also 10 pmol) will be scored as a positive if the integrated optical density (IOD) of the radioactive probe hybridizing to the test DNA (Test) is greater than $IOD_{Neg} + (IOD_{Pos} - IOD_{Neg})/2$. Otherwise, the test DNA will be scored as negative. A positive test correlates with approximately at least 70 % identitiy, and more preferably at least 80-85 sequence identity. If a partial-length open reading frame of the test gene is used, then the equivalent regions of edg-5 and edg-2 will be used as positive and negative controls, respectively, for hybridization.

EXAMPLE 5: Antisense analysis

Knowledge of the correct, complete cDNA sequence of HEDG-5 enables its use as a tool for antisense technology in the investigation of gene function. Oligonucleotides, cDNA or genomic fragments comprising the antisense strand of hedg-5 are used either in vitro or in vivo to inhibit expression of the mRNA. Such technology is now well known in the art, and antisense molecules can be designed at various locations along the nucleotide sequences. By treatment of cells or whole test animals with such antisense sequences, the gene of interest is effectively turned off. Frequently, the function of the gene is ascertained by observing behavior at the intracellular, cellular, tissue or organismal level (e.g., lethality, loss of differentiated function, changes in morphology, etc.).

In addition to using sequences constructed to interrupt transcription of a particular open reading frame, modifications of gene expression is obtained by designing antisense sequences to intron regions, promoter/enhancer elements, or even to trans-acting regulatory genes. Similarly, inhibition is achieved using Hogeboom base-pairing methodology, also known as "triple helix" base pairing.

EXAMPLE 6: Expression of HEDG-5

Expression of hedg-5 is accomplished by subcloning the cDNAs into appropriate expression vectors and transfecting the vectors into analogous expression hosts for example E.Coli. In a particular case, the vector is engineered such that it contains a promoter for β-galactosidase, upstream of the cloning site, followed by sequence containing the amino-terminal Met and the subsequent 7 residues of β-galactosidase. Immediately following these eight residues is an engineered bacteriophage promoter useful for artificial priming and transcription and for providing a number of unique endonuclease restriction sites for cloning.

Induction of the isolated, transfected bacterial strain with IPTG using standard methods produces a fusion protein corresponding to the first seven residues of β-galactosidase, about 15 residues of "linker", and the peptide encoded within the cDNA. Since cDNA clone inserts are generated by an essentially random process, there is one chance in three that the included cDNA will lie in the correct frame for proper translation. If the cDNA is not in the proper reading frame, it is obtained by deletion or insertion of the appropriate number of bases using well known methods including in vitro mutagenesis, digestion with exonuclease III or mung bean nuclease, or the inclusion of an oligonucleotide linker of appropriate length.

The hedg-5 cDNA is shuttled into other vectors known to be useful for expression of protein in specific hosts. Oligonucleotide primers containing cloning sites as well as a segment of DNA (about 25 bases) sufficient to hybridize to stretches at both ends of the target cDNA is synthesized chemically by standard methods. These primers are then used to amplify the desired gene segment by PCR. The resulting gene segment is digested with appropriate restriction enzymes under standard conditions and isolated by gel electrophoresis. Alternately, similar gene segments are produced by digestion of the cDNA with appropriate restriction enzymes. Using appropriate primers, segments of coding sequence from more

than one gene are ligated together and cloned in appropriate vectors. It is possible to optimize expression by construction of such chimeric sequences.

Suitable expression hosts for such chimeric molecules include, but are not limited to, mammalian cells such as Chinese Hamster Ovary (CHO) and human 293 cells, insect cells such as Sf9 cells, yeast cells such as Saccharomyces cerevisiae, and bacteria such as E. coli. For each of these cell systems, a useful expression vector also includes an origin of replication to allow propagation in bacteria and a selectable marker such as the β-lactamase antibiotic resistance gene to allow plasmid selection in bacteria. In addition, the vector may include a second selectable marker such as the neomycin phosphotransferase gene to allow selection in transfected eukaryotic host cells. Vectors for use in eukaryotic expression hosts require RNA processing elements such as 3' polyadenylation sequences if such are not part of the cDNA of interest.

Additionally, the vector contains promoters or enhancers which increase gene expression. Such promoters are host specific and include MMTV, SV40, and metallothionine promoters for CHO cells; trp, lac, tac and T7 promoters for bacterial hosts; and alpha factor, alcohol oxidase and PGH promoters for yeast. Transcription enhancers, such as the rous sarcoma virus enhancer, are used in mammalian host cells. Once homogeneous cultures of recombinant cells are obtained through standard culture methods, large quantities of recombinantly produced HEDG-5 are recovered from the conditioned medium and analyzed using chromatographic methods known in the art. For example, HEDG-5 can be expressibly cloned into the expression vector pcDNA3, as exemplified herein. This product can be used to transform, for example, HEK293 or COS by methodology standard in the art. Specifically, for example, using Lipofectamine (Gibco BRL catalog no. 18324-020) mediated gene transfer.

EXAMPLE 7: Isolation of Recombinant HEDG-5

HEDG-5 is expressed as a chimeric protein with one or more additional polypeptide domains added to facilitate protein purification. Such purification facilitating domains include, but are not limited to, metal chelating peptides such as histidine-tryptophan modules that allow purification on immobilized metals, protein A domains that allow purification on immobilized immunoglobulin, and the domain utilized in the FLAGS extension/affinity

purification system (Immunex Corp., Seattle WA). The inclusion of a cleavable linker sequence such as Factor XA or enterokinase (Invitrogen) between the purification domain and the HEDG-5 sequence is useful to facilitate expression of HEDG-5.

EXAMPLE 8: Testing of Chimeric T7Gs

Functional chimeric T7Gs are constructed by combining the extracellular and/or transmembrane ligand-receptive sequences of a new isoform with the transmembrane and/or intracellular segments of a different T7G for test purposes. This concept was demonstrated by Kobilka et al (1988, Science 240:1310-1316) who created a series of chimeric α2-β2 adrenergic receptors (AR) by inserting progressively greater amounts of α2-AR transmembrane sequence into β2-AR. The binding activity of known agonists changed as the molecule shifted from having more α2 than β2 conformation, and intermediate constructs demonstrated mixed specificity. The specificity for binding antagonists, however, correlated with the source of the domain VII. The importance of T7G domain VII for ligand recognition was also found in chimeras utilizing two yeast α-factor receptors and is significant because the yeast receptors are classified as miscellaneous receptors. Thus, functional role of specific domains appears to be preserved throughout the T7G family regardless of category.

In parallel fashion, internal segments or cytoplasmic domains from a particular isoform are exchanged with the analogous domains of a known T7G and used to identify the structural determinants responsible for coupling the receptors to trimeric G-proteins (Dohlman et al (1991) Annu Rev Biochem 60:653-88). A chimeric receptor in which domains V, VI, and the intracellular connecting loop from β2-AR were substituted into a2-AR was shown to bind ligands with a2-AR specificity, but to stimulate adenylate cyclase in the manner of β2-AR. This demonstrates that for adrenergic-type receptors, G-protein recognition is present in domains V and VI and their connecting loop. The opposite situation was predicted and observed for a chimera in which the V- > VI loop from α1-AR replaced the corresponding domain on β2-AR and the resulting receptor bound ligands with β2-AR specificity and activated G-protein-mediated phosphatidylinositol turnover in the α1-AR manner. Finally, chimeras constructed from muscarinic receptors also demonstrated that V- > VI loop is the major determinant for specificity of G-protein activity (Bolander FF, supra).

Chimeric or modified T7Gs containing substitutions in the extracellular and transmembrane regions have shown that these portions of the receptor determine ligand binding specificity. For example, two Ser residues conserved in domain V of all adrenergic and D catecholamine T7G receptors are necessary for potent agonist activity. These serines are believed to form hydrogen bonds with the catechol moiety of the agonists within the T7G binding site. Similarly, an Asp residue present in domain III of all T7Gs which bind biogenic amines is believed to form an ion pair with the ligand amine group in the T7G binding site.

Functional, cloned T7Gs are expressed in heterologous expression systems and their biological activity assessed (e.g. Marullo et al (1988) Proc Natl Acad Sci 85:7551-55; King et al (1990) Science 250:121-23). One heterologous system introduces genes for a mammalian T7G and a mammalian G-protein into yeast cells. The T7G is shown to have appropriate ligand specificity and affinity and trigger appropriate biological activation--growth arrest and morphological changes--of the yeast cells.

An alternate procedure for testing chimeric receptors is based on the procedure utilizing the P_{2u} purinergic receptor (P_{2u}) as published by Erb et al (1993, Proc Natl Acad Sci 90:104411-53). Function is easily tested in cultured K562 human leukemia cells because these cells lack P_{2u} receptors. K562 cells are transfected with expression vectors containing either normal or chimeric P_{2u} and loaded with fura-a, fluorescent probe for Ca++. Activation of properly assembled and functional P_{2u} receptors with extracellular UTP or ATP mobilizes intracellular Ca++ which reacts with fura-a and is measured spectrofluorometrically. As with the T7G receptors above, chimeric genes are created by combining sequences for extracellular receptive segments of any newly discovered T7G polypeptide with the nucleotides for the transmembrane and intracellular segments of the known P_{2u} molecule. Bathing the transfected K562 cells in microwells containing appropriate ligands triggers binding and fluorescent activity defining effectors of the T7G molecule. Once ligand and function are established, the P_{2u} system is useful for defining antagonists or inhibitors which block binding and prevent such fluorescent reactions.

EXAMPLE 9: Production of HEDG-5 Specific Antibodies

Two approaches are utilized to raise antibodies to HEDG-5, and each approach is useful for generating either polyclonal or monoclonal antibodies. In one approach, denatured

protein from reverse phase HPLC separation is obtained in quantities up to 75 mg. This denatured protein is used to immunize mice or rabbits using standard protocols; about 100 micrograms are adequate for immunization of a mouse, while up to 1 mg might be used to immunize a rabbit. For identifying mouse hybridomas, the denatured protein is radioiodinated and used to screen potential murine B-cell hybridomas for those which produce antibody. This procedure requires only small quantities of protein, such that 20 mg is sufficient for labeling and screening of several thousand clones.

In the second approach, the amino acid sequence of an appropriate HEDG-5 domain, as deduced from translation of the cDNA, is analyzed to determine regions of high antigenicity. Oligopeptides comprising appropriate hydrophilic regions are synthesized and used in suitable immunization protocols to raise antibodies. Analysis to select appropriate epitopes is described by Ausubel FM et al (supra). The optimal amino acid sequences for immunization are usually at the C-terminus, the N-terminus and those intervening, hydrophilic regions of the polypeptide which are likely to be exposed to the external environment when the protein is in its natural conformation.

Typically, selected peptides, about 15 residues in length, are synthesized using an Applied Biosystems Peptide Synthesizer Model 431A using fmoc-chemistry and coupled to keyhole limpet hemocyanin (KLH; Sigma, St. Louis MO) by reaction with M-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS; Ausubel FM et al, supra). If necessary, a cysteine is introduced at the N-terminus of the peptide to permit coupling to KLH. Rabbits are immunized with the peptide-KLH complex in complete Freund's adjuvant. The resulting antisera are tested for antipeptide activity by binding the peptide to plastic, blocking with 1% bovine sewm albumin, reacting with antisera, washing and reacting with labeled (radioactive or fluorescent), affinity purified, specific goat anti-rabbit IgG.

Hybridomas are prepared and screened using standard techniques. Hybridomas of interest are detected by screening with labeled HEDG-5 to identify those fusions producing the monoclonal antibody with the desired specificity. In a typical protocol, wells of plates (FAST; Becton-Dickinson, Palo Alto CA) are coated during incubation with affinity purified, specific rabbit anti-mouse (or suitable antispecies Ig) antibodies at 10 mg/ml. The coated wells are blocked with 1% BSA, washed and incubated with supernatants from hybridomas.

After washing the wells are incubated with labeled HEDG-5 at 1 mg/ml. Supernatants with specific antibodies bind more labeled HEDG-5 than is detectable in the background. Then clones producing specific antibodies are expanded and subjected to two cycles of cloning at limiting dilution. Cloned hybridomas are injected into pristane-treated mice to produce ascites, and monoclonal antibody is purified from mouse ascetic fluid by affinity chromatography on Protein A. Monoclonal antibodies with affinities of at least 10^8 M^{-1}, preferably 10^9 to 10^{10} or stronger, are typically made by standard procedures as described in Harlow and Lane (1988) Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY; and in Goding (1986) Monoclonal Antibodies: Principles and Practice, Academic Press, New York City, both incorporated herein by reference.

EXAMPLE 10: Diagnostic Test Using HEDG-5 Specific Antibodies

Particular HEDG-5 antibodies are useful for investigating signal transduction and the diagnosis of infectious or hereditary conditions which are characterized by differences in the amount or distribution of HEDG-5 or downstream products of an active signaling cascade.

Diagnostic tests for HEDG-5 include methods utilizing antibody and a label to detect HEDG-5 in human body fluids, membranes, cells, tissues or extracts of such. The polypeptides and antibodies of the present invention are used with or without modification. Frequently, the polypeptides and antibodies are labeled by joining them, either covalently or noncovalently, with a substance which provides for a detectable signal. A wide variety of labels and conjugation techniques are known and have been reported extensively in both the scientific and patent literature. Suitable labels include radionuclides, enzymes, substrates, cofactors, inhibitors, fluorescent agents, chemiluminescent agents, chromogenic agents, magnetic particles and the like. Patents teaching the use of such labels include US Patent No's. 3,817,837; 3,850,752; 3,939,350; 3,996,345; 4,277,437; 4,275,149; and 4,366,241. Also, recombinant immunoglobulins may be produced as shown in US Patent No.4,816,567, Incorporated herein by reference.

A variety of protocols for measuring soluble or membrane-bound HEDG-5, using either polyclonal or monoclonal antibodies specific for the protein, are known in the art. Examples include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA) and fluorescent activated cell sorting (FACS). A two-site monoclonal-based immunoassay

utilizing monoclonal antibodies reactive to two non-interfering epitopes on HEDG-5 is preferred, but a competitive binding assay may be employed. These assays are described, among other places, in Maddox, DE et al (1983, J Exp. Med. 158:1211f).

EXAMPLE 11: Purification of Native HEDG-5 Using Specific Antibodies

Native or recombinant HEDG-5 is purified by immunoaffinity chromatography using antibodies specific for HEDG-5. In general, an immunoaffinity column is constructed by covalently coupling the anti-TRH antibody to an activated chromatographic resin.

Polyclonal immunoglobulins are prepared from immune sera either by precipitation with ammonium sulfate or by purification on immobilized Protein A (Pharmacia LKB Biotechnology, Piscataway NJ). Likewise, monoclonal antibodies are prepared from mouse ascites fluid by ammonium sulfate precipitation or chromatography on immobilized Protein A. Partially purified immunoglobulin is covalently attached to a chromatographic resin such as CnBr-activated Sepharose (Pharmacia LKB Biotechnology). The antibody is coupled to the resin, the resin is blocked, and the derivative resin is washed according to the manufacturer's instructions.

Such immunoaffinity columns are utilized in the purification of HEDG-5 by preparing a fraction from cells containing HEDG-5 in a soluble form. This preparation is derived by solubilization of whole cells or of a subcellular fraction obtained via differential centrifugation (with or without addition of detergent) or by other methods well known in the art. Alternatively, soluble HEDG-5 containing a signal sequence is secreted in useful quantity into the medium in which the cells are grown.

A soluble HEDG-5-containing preparation is passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of HEDG-5 (e.g., high ionic strength buffers in the presence of detergent). Then, the column is eluted under conditions that disrupt antibody/protein binding (e.g., a buffer of pH 2-3 or a high concentration of a chaotrope such as urea or thiocyanate ion), and HEDG-5 is collected.

EXAMPLE 12: Drug Screening

This invention is particularly useful for screening therapeutic compounds by using HEDG-5 or binding fragments thereof in any of a variety of drug screening techniques. As HEDG-5 is a G protein coupled receptor any of the methods commonly used in the art may potentially used to identify HEDG-5 ligands. For example, the activity of a G protein coupled receptor such as EDG-5 can be measured using any of a variety of appropriate functional assays in which activation of the receptor results in an observable change in the level of some second messenger system, such as adenylate cyclase, guanylyl cyclase, calcium mobilization, or inositol phospholipid hydrolysis. More particularly, activation of EDG-5 can be measured using the NF-kB, SRE and/or AP-1 functional assays, as described above. One approach, measures the effect of ligand binding on the activation of intracellular second messenger pathways, using a reporter gene. Typically, the reporter gene will have a promoter which is sensitive to the level of that second messenger controlling expression of an easily detectable gene product, for example, CAT or luciferase. Alternatively, the cell is loaded with a reporter substance, e.g., FURA whereby changes in the intracellular concentration of calcium indicate modulation of the receptor as a result of ligand binding. Thus, the present invention provides methods of screening for drugs or any other agents which affect signal transduction.

Alternatively, the polypeptide or fragment employed in such a test is either free in solution, affixed to a solid support, borne on a cell surface or located intracellularly. One method of drug screening utilizes eukaryotic or prokaryotic host cells (or membrane preparations therefrom) which are stably transformed with recombinant nucleic acids expressing the polypeptide or fragment. Drugs are screened against such transformed cells in competition binding assays. ^{32}P-labelled LPA could be used in such a competition binding assay for HEDG-5. Such cells, either in viable or fixed form, are used for standard binding assays. One measures, for example, the formation of complexes between HEDG-5 and the agent being tested. Alternatively, one examines the diminution in complex formation between HEDG-5 and a ligand, for example LPA, caused by the agent being tested.

EXAMPLE 13: Rational Drug Design

Herein, the goal of rational drug design is to produce structural analogs of biologically active phospholipids of interest or of small molecules with which they interact, agonists,

antagonists, or inhibitors. Any of these examples are used to fashion drugs which are more active or stable forms of the phospholipid or which enhance or interfere with the function of a phospholipid in vivo.

In one approach, the three-dimensional structure of a protein of interest, or of a protein-inhibitor complex, is determined by x-ray crystallography, by computer modeling or, most typically, by a combination of the two approaches. Both the shape and charges of the polypeptide must be ascertained to elucidate the structure and to determine active site(s) of the molecule. Less often, useful information regarding the structure of a polypeptide is gained by modeling based on the structure of homologous proteins. In both cases, relevant structural information is used to design efficient inhibitors. Useful examples of rational drug design includes molecules which have improved activity or stability as shown by Braxton S and Wells JA (1992, Biochemistry 31:7796-7801) or which act as inhibitors, agonists, or antagonists of native peptides as shown by Athauda SB et al (1993 J Biochem 113:742-46), incorporated herein by reference.

EXAMPLE 14: Use and Administration of Antibodies, Inhibitors, or Antagonists

Antibodies, inhibitors, or antagonists of HEDG-5 (or other treatments to limit signal transduction, LST) provide different effects when administered therapeutically. LSTs are formulated in a nontoxic, inert, pharmaceutically acceptable aqueous carrier medium preferably at a pH of about 5 to 8, more preferably 6 to 8, although pH may vary according to the characteristics of the antibody, inhibitor, or antagonist being formulated and the condition to be treated. Characteristics of LSTs include solubility of the molecule, half-life and antigenicity/immunogenicity. These and other characteristics aid in defining an effective carrier.

LSTs are delivered by known routes of administration including but not limited to topical creams and gels; transmucosal spray and aerosol; transdermal patch and bandage; injectable, intravenous and lavage formulations; and orally administered liquids and pills particularly formulated to resist stomach acid and enzymes. The particular formulation, exact dosage, and route of administration is determined by the attending physician and varies according to each specific situation.

Such determinations are made by considering multiple variables such as the condition to be treated, the LST to be administered, and the pharmacokinetic profile of a particular LST. Additional factors which are taken into account include severity of the disease state, patient's age, weight, gender and diet, time and frequency of LST administration, possible combination with other drugs, reaction sensitivities, and tolerance/response to therapy. Long acting LST formulations might be administered every 3 to 4 days, every week, or once every two weeks depending on half-life and clearance rate of the particular LST.

Normal dosage amounts vary from 0.1 to 100,000 micrograms, up to a total dose of about 1 g, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature; see US Patent Nos. 4,657,760; 5,206,344; or 5,225,212. Those skilled in the art employ different formulations for different LSTs. Administration to cells such as nerve cells necessitates delivery in a manner different from that to other cells such as vascular endothelial cells.

It is contemplated that abnormal signal transduction, trauma, or diseases which trigger HEDG-5 activity are treatable with LSTs. These conditions or diseases are specifically diagnosed by the tests discussed above, and such testing should be performed in suspected cases of viral, bacterial or fungal infections: allergic responses; mechanical injury associated with trauma; hereditary diseases; lymphoma or carcinoma; or other conditions which activate the genes of lymphoid or neuronal tissues.

EXAMPLE 15: Production of Transgenic Animals

Animal model systems which elucidate the physiological and behavioral roles of the HEDG-5 receptor are produced by creating transgenic animals in which the activity of the HEDG-5 receptor is either increased or decreased, or the amino acid sequence of the expressed HEDG-5 receptor is altered, by a variety of techniques. Examples of these techniques include, but are not limited to: 1) Insertion of normal or mutant versions of DNA encoding a HEDG-5 receptor, by microinjection, electroporation, retroviral transfection or other means well known to those skilled in the art, into appropriate fertilized embryos in order to produce a transgenic animal or 2) Homologous recombination of mutant or normal, human or animal versions of these genes with the native gene locus in transgenic animals to alter the regulation of expression or the structure of these HEDG-5 receptor sequences. The

technique of homologous recombination is well known in the art. It replaces the native gene with the inserted gene and so is useful for producing an animal that cannot express native HEDG-5 receptors but does express, for example, an inserted mutant HEDG-5 receptor, which has replaced the native HEDG-5 receptor in the animal's genome by recombination, resulting in underexpression of the transporter. Microinjection adds genes to the genome, but does not remove them, and so is useful for producing an animal which expresses its own and added HEDG-5 receptors, resulting in overexpression of the HEDG-5 receptors.

One means available for producing a transgenic animal, with a mouse as an example, is as follows: Female mice are mated, and the resulting fertilized eggs are dissected out of their oviducts. The eggs are stored in an appropriate medium such as M2 medium. DNA or cDNA encoding a HEDG-5 purified from a vector by methods well known in the art. Inducible promoters may be fused with the coding region of the DNA to provide an experimental means to regulate expression of the transgene. Alternatively or in addition, tissue specific regulatory elements may be fused with the coding region to permit tissue-specific expression of the trans-gene. The DNA, in an appropriately buffered solution, is put into a microinjection needle (which may be made from capillary tubing using a piper puller) and the egg to be injected is put in a depression slide. The needle is inserted into the pronucleus of the egg, and the DNA solution is injected. The injected egg is then transferred into the oviduct of a pseudopregnant mouse (a mouse stimulated by the appropriate hormones to maintain pregnancy but which is not actually pregnant), where it proceeds to the uterus, implants, and develops to term. As noted above, microinjection is not the only methods for inserting DNA into the egg cell, and is used here only for exemplary purposes.

EXAMPLE 16: Isolation, Chromosomal Localization and Partial Sequencing of a hedg-5 Genomic Clone

To identify genomic clones containing the hedg-5 gene, the H501-20F (SEQ ID NO: 6) and H501-246R (SEQ ID NO: 7) primers were used to amplify human genomic DNA as described in Example 2. One microliter of human genomic DNA (Clontech; Cat #6550-1) was used as template. The PCR product was purified and sequenced in-house, using the PCR primers to prime the sequencing reactions. The sequence of this product (see SEQ ID. NO: 12) matched the cDNA sequence previously obtained for hedg-5 (see SEQ ID. NO: 13),

indicating that these primers could be used to identify genomic clones containing this region of the hedg-5 gene.

An arrayed library of genomic DNA clones (Genome Sciences Inc.) was screened by PCR using these primers. The library contained bacterial artificial chromosome (BAC) constructs with ~120 kb human genomic DNA inserts. In total, clones representing about 3 haploid genome equivalents were screened using the edg-5 diagnostic PCR primers. Two clones were identified by this method: BAC-28 (1F) and BAC-236 (13M). Once the DNA from these clones was received, their identity was verified in-house by sequencing of the PCR product we obtained using the edg-5 diagnostic primers: this analysis showed both clones represent at least part of the hedg-5 gene. The BAC-28 (1F) clone was subsequently used to localize the gene on human chromosomes by fluorescent in situ hybridization (FISH) at Genome Systems Inc. The locus for the hedg-5 gene mapped to band p22.3 of human chromosome 1.

A search of the on-line Mendelian Inheritance in Man database revealed two entries for inherited diseases which genetically map to this region, but for which genes have not yet been cloned. These were the database entries 154280 (Malignant Transformation Suppression-1 or MTS1) and 157900 (Moebius Syndrome). The first represents a dominant suppresser of cellular transformation (a class of genes called tumor suppressers or anti-oncogenes), while the second is an inherited syndrome in which the sixth and seventh cranial nerves are small or absent, leading to facial paralysis. Whether edg-5 gene defects contribute to either of these phenotypes is not known.

Sequencing was performed on DNA prepared from BAC-28 (1F) to determine the position(s) of introns (if any) within the coding region of the edg-5 gene. Sequencing results showed that only one intron exists within the coding region of hedg-5, at a position indicated by the arrowhead between nt 996/997 of the sequence shown in Figure 4A. This intron falls within the codon for Gly-246 of the edg-5 amino acid sequence. Additional sequencing was performed in the region flanking the 5' end of the edg-5 cDNA sequence derived from pC3-hedg-55, revealing 250 bp of genomic DNA sequence upstream of the 5' end of the cDNA.

EXAMPLE 17: Expression and tissue distribution of Edg-5 RNA in the rat.

Northern blotting was carried out with the edg-5 cDNA insert by techniques well-known in the art. Two different multi-tissue rat RNA blots (Origene .Cat. MB-1005 and MB-1007) were probed with radiolabeled edg-5 cDNA. Washing was performed at high stringency conditions that do not permit detection of edg-2 or other related transcripts. The blots were then subjected to autoradiography. The Northern blot results show that RNA expression levels are highest in lung, kidney and testis. Lower RNA levels were seen in skin, heart, small intestine and stomach. Little or no detectable RNA was found in thymus, brain, spleen and liver. Muscle tissue may also express low levels of edg-5 mRNA. Further, anti-sense oligonucletide probes based on the hedg-5 sequence disclosed herein can be used by those of skill in the art to for in situ hybridization expression studies.

Various modifications and variations of the described method and system of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with specific preferred embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments.

CLAIMS

1 An isolated nucleotide sequence encoding a mammalian EDG-5 receptor or biologically active fragment thereof.

2. The isolated nucleotide sequence of claim 1 encoding a murine EDG-5 receptor or biologically active fragment thereof.

3. The isolated nucleotide sequence of claim 2 encoding a murine EDG-5 receptor of Figure 1B or biologically active fragment thereof.

4. The isolated nucleotide sequence of claim 1 encoding a human EDG-5 receptor or biologically active fragment thereof.

5. The isolated nucleotide sequence of claim 1 wherein the biologically active fragment is activated by LPA.

6. An isolated nucleotide sequence selected from the group consisting of:

(a) the nucleotide sequence comprising nucleotides 36-1907 of SEQ. ID NO: 12

(b) the nucleotide sequence of Figure 3B;

(c) the nucleotide sequence of Figure 3C;

(d) a nucleotide sequence comprising at least about 70% sequence identity to (a), (b) or (c) and which hybridizes under stringent conditions to the nucleotide sequence of (a), (b) or (c), respectively; and

(e) the nucleotide sequence which encodes the amino acid sequence of Figure 4A, 4B or 4C.

7. The isolated nucleotide sequence of Claim 6 wherein the nucleotide sequence is selected from the group consisting of:

(1) the nucleotide sequence of (a), (b), (c) or (e) of claim 6; and

(2) the nucleotide sequence of (d) of claim 6 wherein the nucleotide sequence has at least about 80-85% sequence identity to the nucleotide sequence of (a), (b) or (c) of claim 6.

8. The isolated nucleotide sequence of Claim 6 wherein the nucleotide sequence is selected the group consisting of:

(1) the nucleotide sequence of (a), (b), (c) or (e) of claim 6; and

(2) the nucleotide sequence of (d) of claim 6 wherein the nucleotide sequence has at least about 95% sequence identity to the nucleotide sequence of (a), (b) or (c) of claim 6.

9. The complement of the nucleotide sequence of Claim 8.

10. An expression vector comprising the nucleotide sequence of Claim 8.

11. A host cell comprising the expression vector of Claim 10.

12. The isolated and purified amino acid sequence for the HEDG-5 receptor encoded by the nucleotide sequence of claim 8.

13. The isolated and purified amino acid sequence of claim 12 comprising the amino acid sequence of SEQ. ID NO:13 (Figure 4A), Figure 4B or Figure 4C or a biological acitve portion thereof.

14. The isolated nucleotide sequence of Claim 6 wherein the nucleotide sequence is selected from the group consisting of the nucleotide sequence which encodes the amino acid sequence of SEQ ID NO: 13 (Figure 4A), Figure 4B and Figure 4C.

15. A hybridization probe of the nucleotide sequence of Claim 5.

16. A method of screening compounds to identify HEDG-5 ligands comprising the steps of:

(a) culturing cells which express the HEDG-5 receptor or with a membrane preparation obtained therefrom; and

(b) contacting said compound with said cells or said membrane preparation; and

(c) determining whether binding between the HEDG-5 receptor and the candidate ligand has occurred.

17. A HEDG-5 ligand identified by the method of claim 16.

18. A method of screening compounds to identify HEDG-5 antagonists comprising the steps of:

(a) culturing cells which express the HEDG-5 receptor or with a membrane preparation obtained therefrom;

(b) contacting said cells or said membrane preparation with a mixture comprising an agonist and said compound to be tested for antagonist activity at said receptor; and

(3) determining the degree of antagonist activity by measuring a response indicative of the degree of binding between said agonist and said receptor and comparing this measured response with a standard response for binding between said agonist and said receptor absent the antagonist.

19. The method of claim 18 wherein said agonist is LPA.

20. An antagonist identified by the method of claim 18.

ABSTRACT

The present invention is directed to nucleic acid sequences and amino acid sequences for mammalian EDG-5 receptor homologs, and particularly for human EDG-5 receptor homologs. The invention also provides methods for determining agonists and antagonists for EDG-5 receptors in addition to assays, expression vectors, host cells and methods for treating disorders associated with aberrant expression or activity of EDG-5. S1P and SPC are agonists for EDG-5 receptors.

Figure 1A: DNA sequence of the murine edg-5 RT-PCR clone 501 (SEQ ID NO: 3)

```
1    AACACTGGCC CGGTGTCGAA AACGTTGACC GTCAACCGCT GGTTCCTCCG
51   CCAGGGGCTC CTAGACACCA GCCTGACTGC CTCCCTGGCC AATTTGCTGG
101  TTATTGCTGT GGAAAGACAC ATGTCNATCA TGAGGATGAG AGTCCACAGC
151  AACTTGACCA AAAAGCGGGT GACGCTGCTC ATTCTGCTGG TGTGGGCCAT
201  CGCCATCTTC ATGGGGGCCG TCCCCACNCT GGGATGGAAT TGCCTCTGCA
251  ACATCTCGGC CTGCTCTTCT CTGGCTCCCA TTTACAGTAG GAGTTACCTC
301  ATTTTCTGGA CTGTGTCCAA CCTCCTGGCC TTCTTCATCA TGGTGGCGGT
351  ATACGTACGC ATCTACATGT ATGTTAAAAG GAAAACCAAC GTCTTATCTC
401  CACACACCAG TGGCTCCATC AGCCGCCGGA GGGCTCCCAT GAAGCTAATG
451  AAGACAGTGA TGACCGTCTT AGGCGCCTTC GTGGTGTGCT GGACCCCGGG
501  TCTGGTGGTT CTGCTGCTGG ACGGCCTGAA CTGCAAGCAG TGTAACGTGC
551  AACACGTGAA GNGCTGGTTC CTGCTGCTCG CACTGCTCAA CTCCGTCATG
601  AACCCCCTCA TCTACTGCCG CTCTCCNNAC TTTCCATGG
```

Figure 1B: Sequence of full-length mEDG-5 cDNA insert and alignment with MEDG-5 amino acid sequence. Translation starts at nt 19. Translation termination codon is located at nt 1081-1083

```
                       M  N  E  C  H  Y  D  K  R  M  D  F  F  Y
     GCACAGTTCTTGTCCACCATGAATGAGTGTCACTATGACAAGCGCATGGACTTTTTCTAC
1    ---------+---------+---------+---------+---------+---------+  60

      N  R  S  N  T  D  T  A  D  E  W  T  G  T  K  L  V  I  V  L
     AACAGGAGCAACACAGACACAGCGGACGAGTGGACAGGGACAAAGCTTGTGATCGTCCTG
61   ---------+---------+---------+---------+---------+---------+  120

      C  V  G  T  F  F  C  L  F  I  F  F  S  N  S  L  V  I  A  A
     TGCGTGGGGACGTTCTTCTGCCTCTTTATATTTTTTTCTAACTCCCTGGTCATTGCTGCG
121  ---------+---------+---------+---------+---------+---------+  180

      V  I  T  N  R  K  F  H  F  P  F  Y  Y  L  L  A  N  L  A  A
     GTGATCACAAACCGGAAGTTCCACTTTCCCTTCTACTACCTGCTGGCTAACTTAGCTGCT
181  ---------+---------+---------+---------+---------+---------+  240

      A  D  F  F  A  G  I  A  Y  V  F  L  M  F  N  T  G  P  V  S
     GCGGATTTCTTCGCCGGAATCGCTTACGTGTTCCTGATGTTTAACACTGGCCCGGTGTCG
241  ---------+---------+---------+---------+---------+---------+  300

      K  T  L  T  V  N  R  W  F  L  R  Q  G  L  L  D  T  S  L  T
     AAAACGTTGACCGTCAACCGCTGGTTCCTCCGCCAGGGGCTCCTAGACACCAGCCTGACT
301  ---------+---------+---------+---------+---------+---------+  360
```


Figure 1B (cont.)

```
      A  S  L  A  N  L  L  V  I  A  V  E  R  H  M  S  I  M  R  M
     GCCTCCCTGGCCAATTTGCTGGTTATTGCTGTGGAAAGACACATGTCAATCATGAGGATG
361  ---------+---------+---------+---------+---------+---------+  420

      R  V  H  S  N  L  T  K  K  R  V  T  L  L  I  L  L  V  W  A
     AGAGTCCACAGCAACTTGACCAAAAAGCGGGTGACGCTGCTCATTCTGCTGGTGTGGGCC
421  ---------+---------+---------+---------+---------+---------+  480

      I  A  I  F  M  G  A  V  P  T  L  G  W  N  C  L  C  N  I  S
     ATCGCCATCTTCATGGGGGCCGTCCCCACGCTGGGATGGAATTGCCTCTGCAACATCTCG
481  ---------+---------+---------+---------+---------+---------+  540

      A  C  S  S  L  A  P  I  Y  S  R  S  Y  L  I  F  W  T  V  S
     GCCTGCTCTTCTCTGGCTCCCATTTACAGTAGGAGTTACCTCATTTTCTGGACTGTGTCC
541  ---------+---------+---------+---------+---------+---------+  600

      N  L  L  A  F  F  I  M  V  A  V  Y  V  R  I  Y  M  Y  V  K
     AACCTCCTGGCCTTCTTCATCATGGTGGCGGTATACGTACGCATCTACATGTATGTTAAA
601  ---------+---------+---------+---------+---------+---------+  660

      R  K  T  N  V  L  S  P  H  T  S  G  S  I  S  R  R  R  A  P
     AGGAAAACCAACGTCTTATCTCCACACACCAGTGGCTCCATCAGCCGCCGGAGGGCTCCC
661  ---------+---------+---------+---------+---------+---------+  720

      M  K  L  M  K  T  V  M  T  V  L  G  A  F  V  V  C  W  T  P
     ATGAAGCTAATGAAGACAGTGATGACCGTCTTAGGCGCCTTCGTGGTGTGCTGGACCCCG
721  ---------+---------+---------+---------+---------+---------+  780

      G  L  V  V  L  L  L  D  G  L  N  C  K  Q  C  N  V  Q  H  V
     GGTCTGGTGGTTCTGCTGCTGGACGGCCTGAACTGCAAGCAGTGTAACGTGCAACACGTG
781  ---------+---------+---------+---------+---------+---------+  840

      K  R  W  F  L  L  L  A  L  L  N  S  V  M  N  P  I  I  Y  S
     AAGCGCTGGTTCCTGCTGCTCGCACTGCTCAACTCCGTCATGAACCCCATCATCTACTCG
841  ---------+---------+---------+---------+---------+---------+  900

      Y  K  D  E  D  M  Y  N  T  M  R  K  M  I  C  C  A  L  Q  D
     TACAAGGACGAGGACATGTACAACACCATGCGGAAGATGATCTGCTGTGCCCTGCAGGAC
901  ---------+---------+---------+---------+---------+---------+  960

      S  N  T  E  R  R  P  S  R  N  P  S  T  I  H  S  R  S  E  T
     AGCAATACCGAGAGGCGCCCCTCCCGCAACCCCTCCACCATCCACAGCAGGAGCGAGACG
961  ---------+---------+---------+---------+---------+---------+  1020

      G  S  Q  Y  L  E  D  S  I  S  Q  G  P  V  C  N  K  N  G  S
     GGCAGCCAGTACCTGGAGGACAGCATCAGCCAGGGCCCGGTGTGCAATAAAAACGGCTCC
1021 ---------+---------+---------+---------+---------+---------+  1080

      *
     TAAGCCACGGACGCCTCCGCCCTCTTCCCCTGGGGAAAGAGCTGTTAAGCGTCCTCACCT
1081 ---------+---------+---------+---------+---------+---------+  1140

     GTCTCACAAAGCACGTGGACAGGGTTGTTTGAGGGCTCCATGCATCACTTCTGGGGCTTT
1141 ---------+---------+---------+---------+---------+---------+  1200
```

Figure 1B (cont.)

```
        TAAGTTTTCATGGTCAAGGAAAATAGATTTACGGCGTTTAGTAAAGCGCACAGGAAAGGG
1201    ---------+---------+---------+---------+---------+---------+  1260

        AGAGATGAGCAGTGGGTTCCGGCTTGTCTGTGATCCGCTCCCAACATCCTCCAGCTCTTG
1261    ---------+---------+---------+---------+---------+---------+  1320

        CGAGAGCATGCTGGGCTCTGTCACCATCTTGCCACCATTGTCTGTGTGTTTTCAATGATG
1321    ---------+---------+---------+---------+---------+---------+  1380

        GTGTTGAAAGTCCTAGGTCAAAAGAAAGTAGTAAATAATGGTACCTGAGCCCCCCATTGT
1381    ---------+---------+---------+---------+---------+---------+  1440

        GTGGCTACTAGATTCTGTAGTTGTTTCCGCATGGGTTTAAAATGTTCAGAAAAATATTTT
1441    ---------+---------+---------+---------+---------+---------+  1500

        AGCAGTGAACTTTGATTTCCTCAGAGAAGCCATGGCCAGGAGCTAGGTGGGCAACTGTAT
1501    ---------+---------+---------+---------+---------+---------+  1560

        AGTAGAGTAAGTGATGATATTGACCGGTAGGTTGAACTTCTTCCAAATAGCGTCAAATAT
1561    ---------+---------+---------+---------+---------+---------+  1620

        GAGCACGATTAGATCTTCAGTCTTGGTTATCAGGATACCGCTGAGGGGCTTGCTGGATCC
1621    ---------+---------+---------+---------+---------+---------+  1680

        CAAGTGCAAAGTAATTGCACATCGAGTATTTTAACCAAAGCTGCCAGCGTATTCTATCTT
1681    ---------+---------+---------+---------+---------+---------+  1740

        GTGGACTGCATTTTGATCTTGTATTTTTCTCCTTCAAAGACCTCTGAAAGGTAGATCAGT
1741    ---------+---------+---------+---------+---------+---------+  1800

        TAAAAACAAAAATAGTGTTCATACACATAGGCTACTGACCAGTGTTTTCGGTGTAAGACG
1801    ---------+---------+---------+---------+---------+---------+  1860

        TTTAGAGTGTATCTGACAAAGTAAGAATAACTTCAAGGCAGGCACTATGGTATTTATGTA
1861    ---------+---------+---------+---------+---------+---------+  1920

        GCTTGCAAACGTTTACATGTTCTCTCTCTCTCTCTCTCCCCTCTGCTGTTGTGATGTA
1921    ---------+---------+---------+---------+---------+---------+  1980

        ACATTTATGTGCACAAACTACTTGTAATAAAATATTTTAAGAAGCAAAAAAAA
1981    ---------+---------+---------+---------+---------+---  2033
```

Figure 2: Predicted amino acid sequence of Mouse partial EDG-5 cDNA. X represents an amino acid which cannot be assigned due to poor sequencing information from direct PCR sequencing

```
  1  NTGPVSKTLT VNRWFLRQGL LDTSLTASLA NLLVIAVERH MSIMRMRVHS

 51  NLTKKRVTLL ILLVWAIAIF MGAVPTLGWN CLCNISACSS LAPIYSRSYL

101  IFWTVSNLLA FPIMVAVYVR IYMYVKRKTN VLSPHTSGSI SRRRAPMKLM

151  KTVMTVLGAF VVCWTPGLVV LLLDGLNCKQ CNVQHVKXWF LLLALLNSVM

201  NPLIYCRSPX FPW
```


Figure 3A: Nucleotide sequence of a hEDG-5 cDNA inserted into pcDNA3 (SEQ ID NO:13)

Start and stop codons underlined.

```
   1  gaattcgcgg ccgcgtcgac gttcaCTTCT CCACAATGAA TGAGTGTCAC
  51  TATGACAAGC ACATGGACTT TTTTTATAAT AGGAGCAACA CTGATACTGT
 101  CGATGACTGG ACAGGAACAA AGCTTGTGAT TGTTTTGTGT GTTGGGACGT
 151  TTTTCTGCCT GTTTATTTTT TTTTCTAATT CTCTGGTCAT CGCGGCAGTG
 201  ATCAAAAACA GAAAATTTCA TTTCCCCTTT TACTACCTGT TGGCTAATTT
 251  AGCTGCTGCC GATTTCTTCG CTGGAATTGC CTATGTATTC CTGATGTTTA
 301  ACACAGGCCC AGTTTCAAAA ACTTTGACTG TCAACCGCTG GTTTCTCCGT
 351  CAGGGGCTTC TGGACAGTAG CTTGACTGCT TCCCTCACCA ACTTGCTGGT
 401  TATCGCCGTG GAGAGGCACA TGTCAATCAT GAGGATGCGG GTCCATAGCA
 451  ACCTGACCAA AAAGAGGGTG ACACTGCTCA TTTTGCTTGT CTGGGCCATC
 501  GCCATTTTTA TGGGGGCGGT CCCCACACTG GGCTGGAATT GCCTCTGCAA
 551  CATCTCTGCC TGCTCTTCCC TGGCCCCCAT TTACAGCAGG AGTTACCTTG
 601  TTTTCTGGAC AGTGTCCAAC CTCATGGCCT TCCTCATCAT GGTTGTGGTG
 651  TACCTGCGGA TCTACGTGTA CGTCAAGAGG AAAACCAACG TCTTGTCTCC
 701  GCATACAAGT GGGTCCATCA GCCGCCGGAG GACACCCATG AAGCTAATGA
 751  AGACGGTGAT GACTGTCTTA GGGGCGTTTG TGGTATGCTG GACCCCGGGC
 801  CTGGTGGTTC TGCCCCTCGA CGGCCTGAAC TGCAGGCAGT GTGGCGTGCA
 851  GCATGTGAAA AGGTGGTTCC TGCTGCTGGC GCTGCTCAAC TCCGTCGTGA
 901  ACCCCATCAT CTACTCCTAC AAGGACGAGG ACATGTATGG CACCATGAAG
 951  AAGATGATCT GCTGCTTCTC TCAGGAGAAC CCAGAGAGGC GTCCCTCTCG
1001  CATCCCCTCC ACAGTCCTCA GCAGGAGTGA CACAGGCAGC CAGTACATAG
1051  AGGATAGTAT TAGCCAAGGT GCAGTCTGCA ATAAAAGCAC TTCCTAAACT
1101  CTGGATGCCT CTYGGCCCAC CCAGGCCTCC TCTGGGAAAA GAGCTGTTAA
1151  GAATGATTAC CTGTCTCTAA CAAAGCCCAT GTACAGTGTT ATTTGAGGTC
1201  TCCATTAATC ACTGCTAGAT TTCTTTAAAA AATTTTTTTT CATAGTTTAA
1251  AAGCATGGGC AGTAAAGAGA GGACCTGCTG CATTTAGAGA AAGCACAGgt
1301  cgacgcggcc gcgaattctt ttgcttttta ccctggaaga aatactcgag
1351  catgcat
```

Figure 3B: cDNA sequence of clone pC3-hEDG5#3.4 from the region encoding a hEDG5 polypeptide

```
   1 ATGAATGAGT GTCACTATGA CAAGCACATG GACTTTTTTT ATAATAGGAG
  51 CAACACTGAT ACTGTCGATG ACTGGACTGG AACAAAGCTT GTGATTGTTT
 101 TGTGTGTTGG GACGTTTTTC TGCCTGTTTA TTTTTTTTTC TAATTCTCTG
 151 GTCATCGCGG CAGTGATCAA AAACAGAAAA TTTCATTTCC CCTTCTACTA
 201 CCTGTTGGCT AATTTGGCTG CTGCCGATTT CTTCGCTGGA ATTGCCTATG
 251 TATTCCTGAT GTTTAACACA GGCCCAGTTT CAAAAACTTT GACTGTCAAC
 301 CGCTGGTTTC TCCGTCAGGG GCTTCTGGAC AGTAGCTTGA CTGCTTCCCT
 351 CACCAACTTG CTGGTTATCG CCGTGGAGAG GCACATGTCA ATCATGAgGA
 401 TGCGGGTCCA TAGCAACCTG ACCAAAAAGA GGGTGACACT GCTCATTTTG
 451 CTTGTCTGGG CCATCGCCAT TTTTATGGGG GCGGTCCCCA CACTGGGCTG
 501 GAATTGCCTC TGCAAcATCT CTGCCTGcTC TTCCCTGGCC CCCATTTACA
 551 GCAGGAGTTA CCTTGTTTTc TGGACAGTGT CCAACCTCAT GGCCTTCCTc
 601 ATCATGGTTG TGGTGTACCT GCGGATCTAC GTGTACGTCA AGAGGAAAAC
 651 CAACGTCTTG TCTCCGCATA CAAGTGGGTC CATCAGCCGC CGGAGGACAC
 701 CCATGAAGCT AATGAAGACG GTAATGACTG TCTTAGGGGC GTTTGTGGTA
 751 TGCTGGACCC CGGGCCTGGT GGTTCTGCTC CTCGACGGCC TGAACTGCAG
 801 GCAGTGTGGC GTGCAGCATG TGAAAAGGTG GTTCCTGCTG CTGGCGCTGC
 851 TCAACTCCGT CGTGAACCCT ATCATCTACT CCTACAAGGA CGAGGACATG
 901 TATGGCACCA TGAAGAAGAT GATCTGCTGC TTCTCTCAGG AGAACCCAGA
 951 GAGGCGTCCC TCTCGCATCC CCTCCACAGT CCTCAGCAGG AGTGACACAG
1001 GCAGCCAGTA CATAGAGGAT AGTATTAGCC AAGGTGCAGT CTGCAATAAA
1051 AGCACTTCCT AA
```


Figure 3C: cDNA sequence of clone pC3-hEDG5#28 from the region encoding a hEDG5 polypeptide.

```
   1 ATGAATGAGT GTCACTATGA CAAGCACATG GACTTTTTTT ATAATAGGGG
  51 CAACACTGAT ACTGTCGATG ACTGGACAGG AACAAAGCTT GTGATTGTTT
 101 TGTGTGTtGG GACGTTTTTC TGCCTGTTTA TTTTTTTTTC TAATTCTCTG
 151 GTCATCGCGG CAGTGATCAA AAACAGAAAA TTTCATTTCC CCTTCTACTA
 201 CCTGTTGGCT AATTTAGCTG CTGCCGATTT CTTCGCTGGA ATTGCCTATG
 251 TATTCCTGAT GTTTAACACA GGCCCAGTTT CAAAAACTTT GACTGTCAAC
 301 CGCTGGTTTC TCCGTCAGGG GCTTCTGGAC AGTAGCTTGA CTGCTTCCCT
 351 CACCAACTTG CTGGTTATCG CCGTGGAGAG GCACATGTCA ATCATGAGGA
 401 TGCGGGTCCA TAGCAACCTG ACCAAAAAGA GGGTGACACT GCTCATTTTG
 451 CTTGTCTGGG CCATCGCCAT TTTTATGGGG GCGGTCCCCA CACTGGGCtG
 501 GAATTGCCTC TGCAAcATcT CTGCCTGcTC TTCCCTGGCC CCCATTTACA
 551 GCAGGAGTTA CCTTGTTTTC TGGACAGTGT CcAACCTCAT GGCcTTCCTC
 601 ATCATGGTTG TGGTGTACCT GCGGATCtAC GTGtACGTCA AGAGGAAAAC
 651 CAAcGTCTTG TCTCCGCAtA CAAGTGGGTC CATCAGCCGC CGGAGGACAC
 701 CCATGAAGCT AATGAAGACG GTGATGACTG TCTTAGGGGC GTTTGTGGTA
 751 TGCTGGACCC CGGGCCTGGT GGTTCTGCTC CTCGACGGCC TGAACTGCAG
 801 GCAGTGTGGC GTGCAGCATG TGAAAAGGTG GTTCCTGCTG CTGGCGCTGC
 851 TCAACTCCGT CGTGAACCCC ATCATCTACT CCTACAAGGA CGAGGACATG
 901 TATGGCACCA TGAAGAAgAT GATCTGCTGC TTCTCTCAGG AGAACCCAGA
 951 GAGGCGTCCC TCTCGCATCC CCTCCACAGT CCTCAGCAGG AGTGACACAG
1001 GCAGCCAGTA CATAGAGGAT AGTATTAGCC AAGGTGCAGT CTGCAATAAA
1051 AGCACTTCCT AA
```

Figure 4A: Aligned hedg-5 cDNA and predicted amino acid sequence. The first 250 bp of DNA sequence (lower case) is derived from genomic DNA flanking the 5' end of the cDNA insert from clone pC3-hedg-55. Sequences from nt 251-1523 are shown in lower case wherever apparent polymorphisms in different human clones were found. Coding region polymorphisms are detailed in Table 1. One intron exists within the coding region of hedg-5, located between nt 996/997 of the cDNA sequence shown.

```
      caccttcctaacctgagcggcctagcctgggaaacaaacaattaaaatgtgcgctaaatg
   1  ---------+---------+---------+---------+---------+---------+  60
      gtggaaggattggactcgccggatcggaccctttgtttgttaattttacacgcgatttac

      ctgtggtaggaggtcaggggctatgtcctggaccaaaggacatttgcactgagacctgac
  61  ---------+---------+---------+---------+---------+---------+  120
      gacaccatcctccagtccccgatacaggacctggtttcctgtaaacgtgactctggactg

      acttcaggtcttcaactcccttgatgggagttagccagaacgggcttagaaacagcaatt
 121  ---------+---------+---------+---------+---------+---------+  180
      tgaagtccagaagttgagggaactaccctcaatcggtcttgcccgaatctttgtcgttaa

      gatggcttagtgactgattttacaaatgatatttgtttcttctttaaatttctttctagg
 181  ---------+---------+---------+---------+---------+---------+  240
      ctaccgaatcactgactaaaatgtttactataaacaaagaagaaatttaaagaaagatcc

                              M  N  E  C  H  Y  D  K  H  M  D  F  F  Y
      atgttcacttCTTCTCCACAATGAATGAGTGTCACTATGACAAGCACATGGACTTTTTTT
 241  ---------+---------+---------+---------+---------+---------+  300
      tacaagtgaagaagAGGTGTTACTTACTCACAGTGATACTGTTCGTGTACCTGAAAAAAA

       N  R  S  N  T  D  T  V  D  D  W  T  G  T  K  L  V  I  V  L
      ATAATAGGAGCAACACTGATACTGTCGATGACTGGACAGGAACAAAGCTTGTGATTGTTT
 301  ---------+---------+---------+---------+---------+---------+  360
      TATTATCCTCGTTGTGACTATGACAGCTACTGACCTGTCCTTGTTTCGAACACTAACAAA

       C  V  G  T  F  F  C  L  F  I  F  F  S  N  S  L  V  I  A  A
      TGTGTGTTGGGACGTTTTTCTGCCTGTTTATTTTTTTTTCTAATTCTCTGGTCATCGCGG
 361  ---------+---------+---------+---------+---------+---------+  420
      ACACACAACCCTGCAAAAAGACGGACAAATAAAAAAAAAGATTAAGAGACCAGTAGCGCC

       V  I  K  N  R  K  F  H  F  P  F  Y  Y  L  L  A  N  L  A  A
      CAGTGATCAAAAACAGAAAATTTCATTTCCCCTTTTACTACCTGTTGGCTAATTTAGCTG
 421  ---------+---------+---------+---------+---------+---------+  480
      GTCACTAGTTTTTGTCTTTTAAAGTAAAGGGGAAAATGATGGACAACCGATTAAATCGAC

       A  D  F  F  A  G  I  A  Y  V  F  L  M  F  N  T  G  P  V  S
      CTGCCGATTTCTTCGCTGGAATTGCCTATGTATTCCTGATGTTTAACACAGGCCCAGTTT
 481  ---------+---------+---------+---------+---------+---------+  540
      GACGGCTAAAGAAGCGACCTTAACGGATACATAAGGACTACAAATTGTGTCCGGGTCAAA

       K  T  L  T  V  N  R  W  F  L  R  Q  G  L  L  D  S  S  L  T
      CAAAAACTTTGACTGTCAACCGCTGGTTTCTCCGTCAGGGGCTTCTGGACAGTAGCTTGA
 541  ---------+---------+---------+---------+---------+---------+  600
      GTTTTTGAAACTGACAGTTGGCGACCAAAGAGGCAGTCCCCGAAGACCTGTCATCGAACT
```


Figure 4A (cont.)

```
       A  S  L  T  N  L  L  V  I  A  V  E  R  H  M  S  I  M  R  M
      CTGCTTCCCTCACCAACTTGCTGGTTATCGCCGTGGAGAGGCACATGTCAATCATGAGGA
 601  ---------+---------+---------+---------+---------+---------+  660
      GACGAAGGGAGTGGTTGAACGACCAATAGCGGCACCTCTCCGTGTACAGTTAGTACTCCT

       R  V  H  S  N  L  T  K  K  R  V  T  L  L  I  L  L  V  W  A
      TGCGGGTCCATAGCAACCTGACCAAAAAGAGGGTGACACTGCTCATTTTGCTTGTCTGGG
 661  ---------+---------+---------+---------+---------+---------+  720
      ACGCCCAGGTATCGTTGGACTGGTTTTTCTCCCACTGTGACGAGTAAAACGAACAGACCC

       I  A  I  F  M  G  A  V  P  T  L  G  W  N  C  L  C  N  I  S
      CCATCGCCATTTTTATGGGGGCGGTCCCCACACTGGGCTGGAATTGCCTCTGCAACATCT
 721  ---------+---------+---------+---------+---------+---------+  780
      GGTAGCGGTAAAAATACCCCCGCCAGGGGTGTGACCCGACCTTAACGGAGACGTTGTAGA

       A  C  S  S  L  A  P  I  Y  S  R  S  Y  L  V  F  W  T  V  S
      CTGCCTGCTCTTCCCTGGCCCCCATTTACAGCAGGAGTTACCTTGTTTTCTGGACAGTGT
 781  ---------+---------+---------+---------+---------+---------+  840
      GACGGACGAGAAGGGACCGGGGGTAAATGTCGTCCTCAATGGAACAAAAGACCTGTCACA

       N  L  M  A  F  L  I  M  V  V  V  Y  L  R  I  Y  V  Y  V  K
      CCAACCTCATGGCCTTCCTCATCATGGTTGTGGTGTACCTGCGGATCTACGTGTACGTCA
 841  ---------+---------+---------+---------+---------+---------+  900
      GGTTGGAGTACCGGAAGGAGTAGTACCAACACCACATGGACGCCTAGATGCACATGCAGT

       R  K  T  N  V  L  S  P  H  T  S  G  S  I  S  R  R  R  T  P
      AGAGGAAAACCAACGTCTTGTCTCCGCATACAAGTGGGTCCATCAGCCGCCGGAGGACAC
 901  ---------+---------+---------+---------+---------+---------+  960
      TCTCCTTTTGGTTGCAGAACAGAGGCGTATGTTCACCCAGGTAGTCGGCGGCCTCCTGTG

       M  K  L  M  K  T  V  M  T  V  L  G  A  F  V  V  C  W  T  P
      CCATGAAGCTAATGAAGACGGTGATGACTGTCTTAGGGGCGTTTGTGGTATGCTGGACCC
 961  ---------+---------+---------+---------+---------+---------+  1020
      GGTACTTCGATTACTTCTGCCACTACTGACAGAATCCCCGCAAACACCATACGACCTGGG

       G  L  V  V  L  P  L  D  G  L  N  C  R  Q  C  G  V  Q  H  V
      CGGGCCTGGTGGTTCTGCCCCTCGACGGCCTGAACTGCAGGCAGTGTGGCGTGCAGCATG
1021  ---------+---------+---------+---------+---------+---------+  1080
      GCCCGGACCACCAAGACGGGGAGCTGCCGGACTTGACGTCCGTCACACCGCACGTCGTAC

       K  R  W  F  L  L  L  A  L  L  N  S  V  V  N  P  I  I  Y  S
      TGAAAAGGTGGTTCCTGCTGCTGGCGCTGCTCAACTCCGTCGTGAACCCCATCATCTACT
1081  ---------+---------+---------+---------+---------+---------+  1140
      ACTTTTCCACCAAGGACGACGACCGCGACGAGTTGAGGCAGCACTTGGGGTAGTAGATGA

       Y  K  D  E  D  M  Y  G  T  M  K  K  M  I  C  C  F  S  Q  E
      CCTACAAGGACGAGGACATGTATGGCACCATGAAGAAGATGATCTGCTGCTTCTCTCAGG
1141  ---------+---------+---------+---------+---------+---------+  1200
      GGATGTTCCTGCTCCTGTACATACCGTGGTACTTCTTCTACTAGACGACGAAGAGAGTCC

       N  P  E  R  R  P  S  R  I  P  S  T  V  L  S  R  S  D  T  G
      AGAACCCAGAGAGGCGTCCCTCTCGCATCCCCTCCACAGTCCTCAGCAGGAGTGACACAG
1201  ---------+---------+---------+---------+---------+---------+  1260
      TCTTGGGTCTCTCCGCAGGGAGAGCGTAGGGGAGGTGTCAGGAGTCGTCCTCACTGTGTC

       S  Q  Y  I  E  D  S  I  S  Q  G  A  V  C  N  K  S  T  S  *
      GCAGCCAGTACATAGAGGATAGTATTAGCCAAGGTGCAGTCTGCAATAAAAGCACTTCCT
1261  ---------+---------+---------+---------+---------+---------+  1320
      CGTCGGTCATGTATCTCCTATCATAATCGGTTCCACGTCAGACGTTATTTTCGTGAAGGA
```

Figure 4A (cont.)

```
      AAACTCTGGATGCCTCTYGGCCCACCCAGGCCTCCTCTGGGAAAAGAGCTGTTAAGAATG
1321  ---------+---------+---------+---------+---------+---------+  1380
      TTTGAGACCTACGGAGARCCGGGTGGGTCCGGAGGAGACCCTTTTCTCGACAATTCTTAC

      ATTACCTGTCTCTAACAAAGCCCATGTACAGTGTTATTTGAGGTCTCCATTAATCACTGC
1381  ---------+---------+---------+---------+---------+---------+  1440
      TAATGGACAGAGATTGTTTCGGGTACATGTCACAATAAACTCCAGAGGTAATTAGTGACG

      TAGATTTCTTTAAAAAATTTTTTTTCATAGTTTAAAAGCATGGGCAGTAAAGAGAGGACC
1441  ---------+---------+---------+---------+---------+---------+  1500
      ATCTAAAGAAATTTTTTAAAAAAAAGTATCAAATTTTCGTACCCGTCATTTCTCTCCTGG

      TGCTGCATTTAGAGAAAGCACAG
1501  ---------+---------+---  1523
      ACGACGTAAATCTCTTTCGTGTC
```

Figure 4B: Predicted amino acid sequence of hEDG5 encoded by clone pC3-hEDG5#3.4

```
  1  MNECHYDKHM DFFYNRSNTD TVDDWTGTKL VIVLCVGTFF CLFIFFSNSL
 51  VIAAVIKNRK FHFPFYYLLA NLAAADFFAG IAYVFLMFNT GPVSKTLTVN
101  RWFLRQGLLD SSLTASLTNL LVIAVERHMS IMRMRVHSNL TKKRVTLLIL
151  LVWAIAIFMG AVPTLGWNCL CNISACSSLA PIYSRSYLVF WTVSNLMAFL
201  IMVVVYLRIY VYVKRKTNVL SPHTSGSISR RRTPMKLMKT VMTVLGAFVV
251  CWTPGLVVLL LDGLNCRQCG VQHVKRWFLL LALLNSVVNP IIYSYKDEDM
301  YGTMKKMICC FSQENPERRP SRIPSTVLSR SDTGSQYIED SISQGAVCNK
351  STS
```

Figure 4C: Predicted amino acid sequence of hEDG5 encoded by clone pC3-hEDG5#28.

```
  1  MNECHYDKHM DFFYNRGNTD TVDDWTGTKL VIVLCVGTFF CLFIFFSNSL
 51  VIAAVIKNRK FHFPFYYLLA NLAAADFFAG IAYVFLMFNT GPVSKTLTVN
101  RWFLRQGLLD SSLTASLTNL LVIAVERHMS IMRMRVHSNL TKKRVTLLIL
151  LVWAIAIFMG AVPTLGWNCL CNISACSSLA PIYSRSYLVF WTVSNLMAFL
201  IMVVVYLRIY VYVKRKTNVL SPHTSGSISR RRTPMKLMKT VMTVLGAFVV
251  CWTPGLVVLL LDGLNCRQCG VQHVKRWFLL LALLNSVVNP IIYSYKDEDM
301  YGTMKKMICC FSQENPERRP SRIPSTVLSR SDTGSQYIED SISQGAVCNK
351  STS
```


Figure 5A: Alignment of predicted amino acid sequences of HEDG5 translation products of clones pC3-hEdg5#3.4, pC3-hEdg5-3 and pC3-hEdg5#28. Amino acid substitutions are indicated in reverse bold text.

```
                 1                                                  50
he5#3.4_981216   MNECHYDKHM DFFYNRSNTD TVDDWTGTKL VIVLCVGTFF CLFIFFSNSL
          hsedg5 MNECHYDKHM DFFYNRSNTD TVDDWTGTKL VIVLCVGTFF CLFIFFSNSL
   he5#28_981215 MNECHYDKHM DFFYNRGNTD TVDDWTGTKL VIVLCVGTFF CLFIFFSNSL

                 51                                                100
he5#3.4_981216   VIAAVIKNRK FHFPFYYLLA NLAAADFFAG IAYVFLMFNT GPVSKTLTVN
          hsedg5 VIAAVIKNRK FHFPFYYLLA NLAAADFFAG IAYVFLMFNT GPVSKTLTVN
   he5#28_981215 VIAAVIKNRK FHFPFYYLLA NLAAADFFAG IAYVFLMFNT GPVSKTLTVN

                 101                                               150
he5#3.4_981216   RWFLRQGLLD SSLTASLTNL LVIAVERHMS IMRMRVHSNL TKKRVTLLIL
          hsedg5 RWFLRQGLLD SSLTASLTNL LVIAVERHMS IMRMRVHSNL TKKRVTLLIL
   he5#28_981215 RWFLRQGLLD SSLTASLTNL LVIAVERHMS IMRMRVHSNL TKKRVTLLIL

                 151                                               200
he5#3.4_981216   LVWAIAIFMG AVPTLGWNCL CNISACSSLA PIYSRSYLVF WTVSNLMAFL
          hsedg5 LVWAIAIFMG AVPTLGWNCL CNISACSSLA PIYSRSYLVF WTVSNLMAFL
   he5#28_981215 LVWAIAIFMG AVPTLGWNCL CNISACSSLA PIYSRSYLVF WTVSNLMAFL

                 201                                               250
he5#3.4_981216   IMVVVYLRIY VYVKRKTNVL SPHTSGSISR RRTPMKLMKT VMTVLGAFVV
          hsedg5 IMVVVYLRIY VYVKRKTNVL SPHTSGSISR RRTPMKLMKT VMTVLGAFVV
   he5#28_981215 IMVVVYLRIY VYVKRKTNVL SPHTSGSISR RRTPMKLMKT VMTVLGAFVV

                 251                                               300
he5#3.4_981216   CWTPGLVVLL LDGLNCRQCG VQHVKRWFLL LALLNSVVNP IIYSYKDEDM
          hsedg5 CWTPGLVVLP LDGLNCRQCG VQHVKRWFLL LALLNSVVNP IIYSYKDEDM
   he5#28_981215 CWTPGLVVLL LDGLNCRQCG VQHVKRWFLL LALLNSVVNP IIYSYKDEDM

                 301                                               350
he5#3.4_981216   YGTMKKMICC FSQENPERRP SRIPSTVLSR SDTGSQYIED SISQGAVCNK
          hsedg5 YGTMKKMICC FSQENPERRP SRIPSTVLSR SDTGSQYIED SISQGAVCNK
   he5#28_981215 YGTMKKMICC FSQENPERRP SRIPSTVLSR SDTGSQYIED SISQGAVCNK

                 351
he5#3.4_981216   STS-
          hsedg5 STS-
   he5#28_981215 STS-
```

Figure 5B: Alignment of the amino acid sequence of the murine edg-5 with the amino acid sequence of the human edg-5 of the pC3-hEdg5-3.4 clone.

```
SCORES      Init1:  1981  Initn:  2155  Opt:  2170 z-score: 384.5 E():  2e-16
Smith-Waterman score: 2170;    91.2% identity in 354 aa overlap

                  10        20        30        40        50        60
MEDG5      MNECHYDKRMDFFYNRSNTDTADEWTGTKLVIVLCVGTFFCLFIFFSNSLVIAAVITNRK
           ||||||||:||||||||||||:|:|||||||||||||||||||||||||||||||| |||
HEDG5-3.4  MNECHYDKHMDFFYNRSNTDTVDDWTGTKLVIVLCVGTFFCLFIFFSNSLVIAAVIKNRK
                  10        20        30        40        50        60

                  70        80        90       100       110       120
MEDG5      FHFPFYYLLANLAAADFFAGIAYVFLMFNTGPVSKTLTVNRWFLRQGLLDTSLTASLANL
           ||||||||||||||||||||||||||||||||||||||||||||||||||:||||||:||
HEDG5-3.4  FHFPFYYLLANLAAADFFAGIAYVFLMFNTGPVSKTLTVNRWFLRQGLLDSSLTASLTNL
                  70        80        90       100       110       120

                 130       140       150       160       170       180
MEDG5      LVIAVERHMSIMRMRVHSNLTKKRVTLLILLVWAIAIFMGAVPTLGWNCLCNISACSSLA
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
HEDG5-3.4  LVIAVERHMSIMRMRVHSNLTKKRVTLLILLVWAIAIFMGAVPTLGWNCLCNISACSSLA
                 130       140       150       160       170       180

                 190       200       210       220       230       240
MEDG5      PIYSRSYLIFWTVSNLLAFFIMVAVYVRIYMYVKRKTNVLSPHTSGSISRRRAPMKLMKT
           ||||||||:|||||||:||:|||:||:|||:|||||||||||||||||||||:|||||||
HEDG5-3.4  PIYSRSYLVFWTVSNLMAFLIMVVVYLRIYVYVKRKTNVLSPHTSGSISRRRTPMKLMKT
                 190       200       210       220       230       240

                 250       260       270       280       290       300
MEDG5      VMTVLGAFVVCWTPGLVVLLLDGLNCKQCNVQHVKRWFLLLALLNSVMNPIIYSYKDEDM
           ||||||||||||||||||||||||||:||.|||||||||||||||||:||||||||||||
HEDG5-3-4  VMTVLGAFVVCWTPGLVVLLLDGLNCRQCGVQHVKRWFLLLALLNSVVNPIIYSYKDEDM
                 250       260       270       280       290       300

                 310       320       330       340       350
MEDG5      YNTHRKMICCALQDSNTERRPSRNPSTIHSRSETGSQYLEDSISQGPVCNKNGS
           |:||:|||||  |: | |||||| |||: |||:||||:||||||| ||||: |
HEDG5-3.4  YGTMKKMICCPFSQE-NPERRPSRIPSTVLSRSDTGSQYIEDSISQGAVCNKSTS
                 310       320       330       340       350
```


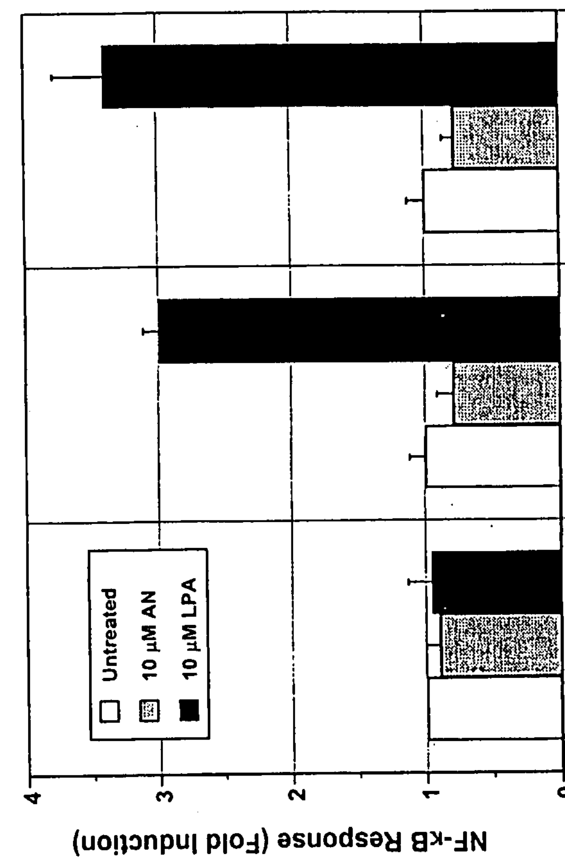

Figure 6

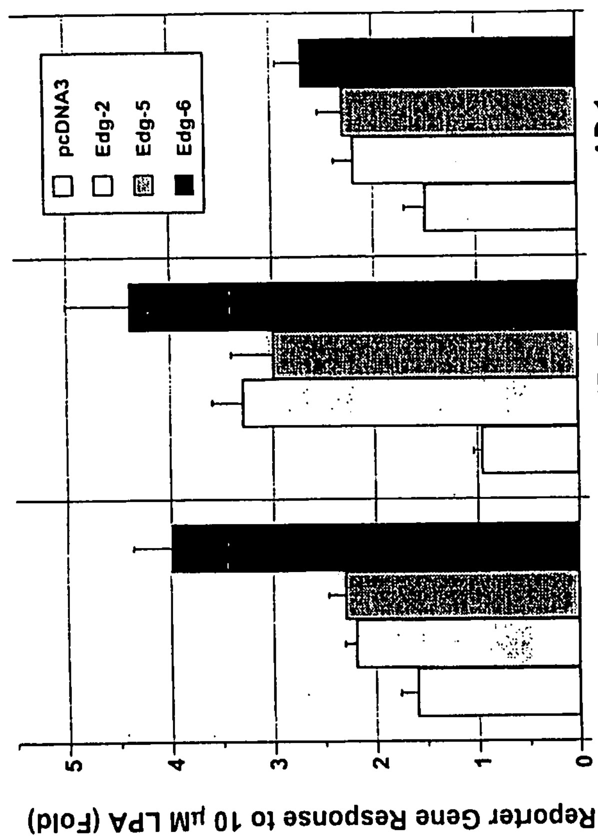

Figure 7

SEQUENCE LISTINGS

(1) GENERAL INFORMATION:

(i) APPLICANT:
NAME: ALLELIX BIOPHARMACEUTICALS INC.
STREET: 6850 Goreway Drive
CITY: Mississauga
PROVINCE: Ontario
COUNTRY: Canada
POSTAL CODE: L4V 1V7
TELEPHONE: (905) 677-0831
FACSIMILE: (905) 677-9595

(ii) TITLE OF INVENTION: MAMMALIAN EDG-5 RECEPTOR HOMOLOGS

(iii) NUMBER OF SEQUENCES: 15

(iv) CORRESPONDENCE ADDRESS:
(A) NAME: Orange Chari Pillay
(B) STREET: Suite 3600, P.O. Box 190
Toronto Dominion Bank Tower
Toronto-Dominion Centre
(C) CITY: Toronto
(D) PROVINCE: Ontario
(E) COUNTRY: Canada
(F) ZIP: M5K 1H6

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: DOS EDITOR

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: PCT/CA98/01193
(B) FILING DATE: 24-DEC-1998
(C) CLASSIFICATION:

(vi) PRIOR APPLICATION DATA:
(A) COUNTRY: U.S.A.
(B) APPLICATION NUMBER: 08/997,803
(C) FILING DATE: 24-DEC-1997
(D) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Santosh K. Chari
(B) FIRM: Orange Chari Pillay
(C) REFERENCE NUMBER: 8700213-0006
(D) TELEPHONE: (416) 868-3457
(E) FACSIMILE: (416) 364-7910

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
AAYTRSATMT STAAYYTGCG TGCGA                                   25
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:
(A) NAME/KEY: Modified Site
(B) LOCATION:4
(D) OTHER INFORMATION: N is Inosine

(ix) FEATURE:
(A) NAME/KEY: Modified Site
(B) LOCATION:13
(D) OTHER INFORMATION: N is Inosine

(ix) FEATURE:
(A) NAME/KEY: Modified Site
(B) LOCATION:16
(D) OTHER INFORMATION: N is Inosine

(ix) FEATURE:
(A) NAME/KEY: Modified Site
(B) LOCATION:22
(D) OTHER INFORMATION: N is Inosine

(ix) FEATURE:
(A) NAME/KEY: Modified Site
(B) LOCATION:25
(D) OTHER INFORMATION: N is Inosine

(ix) FEATURE:
(A) NAME/KEY: Modified Site
(B) LOCATION:28
(D) OTHER INFORMATION: N is Inosine

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
CTGNYKWTTC ATNAWNMMRT ANAYNAYNGG RTT                          33
```


(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 639 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..634

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```
AAC ACT GGC CCG GTG TCG AAA ACG TTG ACC GTC AAC CGC TGG TTC CTC      48
Asn Thr Gly Pro Val Ser Lys Thr Leu Thr Val Asn Arg Trp Phe Leu
  1               5                  10                  15

CGC CAG GGG CTC CTA GAC ACC AGC CTG ACT GCC TCC CTG GCC AAT TTG      96
Arg Gln Gly Leu Leu Asp Thr Ser Leu Thr Ala Ser Leu Ala Asn Leu
             20                  25                  30

CTG GTT ATT GCT GTG GAA AGA CAC ATG TCN ATC ATG AGG ATG AGA GTC     144
Leu Val Ile Ala Val Glu Arg His Met Ser Ile Met Arg Met Arg Val
         35                  40                  45

CAC AGC AAC TTG ACC AAA AAG CGG GTG ACG CTG CTC ATT CTG CTG GTG     192
His Ser Asn Leu Thr Lys Lys Arg Val Thr Leu Leu Ile Leu Leu Val
     50                  55                  60

TGG GCC ATC GCC ATC TTC ATG GGG GCC GTC CCC ACN CTG GGA TGG AAT     240
Trp Ala Ile Ala Ile Phe Met Gly Ala Val Pro Thr Leu Gly Trp Asn
 65                  70                  75                  80

TGC CTC TGC AAC ATC TCG GCC TGC TCT TCT CTG GCT CCC ATT TAC AGT     288
Cys Leu Cys Asn Ile Ser Ala Cys Ser Ser Leu Ala Pro Ile Tyr Ser
                 85                  90                  95
AGG AGT TAC CTC ATT TTC TGG ACT GTG TCC AAC CTC CTG GCC TTC TTC     336
Arg Ser Tyr Leu Ile Phe Trp Thr Val Ser Asn Leu Leu Ala Phe Phe
            100                 105                 110

ATC ATG GTG GCG GTA TAC GTA CGC ATC TAC ATG TAT GTT AAA AGG AAA     384
Ile Met Val Ala Val Tyr Val Arg Ile Tyr Met Tyr Val Lys Arg Lys
        115                 120                 125

ACC AAC GTC TTA TCT CCA CAC ACC AGT GGC TCC ATC AGC CGC CGG AGG     432
Thr Asn Val Leu Ser Pro His Thr Ser Gly Ser Ile Ser Arg Arg Arg
    130                 135                 140

GCT CCC ATG AAG CTA ATG AAG ACA GTG ATG ACC GTC TTA GGC GCC TTC     480
Ala Pro Met Lys Leu Met Lys Thr Val Met Thr Val Leu Gly Ala Phe
145                 150                 155                 160

GTG GTG TGC TGG ACC CCG GGT CTG GTG GTT CTC CTG CTG GAC GGC CTG     528
Val Val Cys Trp Thr Pro Gly Leu Val Val Leu Leu Leu Asp Gly Leu
                165                 170                 175
```

```
AAC TGC AAG CAG TGT AAC GTG CAA CAC GTG AAG NGC TGG TTC CTG CTG       576
Asn Cys Lys Gln Cys Asn Val Gln His Val Lys Xaa Trp Phe Leu Leu
        180                 185                 190

CTC GCA CTG CTC AAC TCC GTC ATG AAC CCC CTC ATC TAC TGC CGC TCT       624
Leu Ala Leu Leu Asn Ser Val Met Asn Pro Leu Ile Tyr Cys Arg Ser
        195                 200                 205

CCN NAC TTT CCA TGG                                                   639
Pro Xaa Phe Pro Trp
    210
```

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 37 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

TTTTTACTCG AGATTTGCTG GTTATTGCTG TGGAAAG 37

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 37 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

TTTTTTCTAG ACGGTCATCA CTGTCTTCAT TAGCTTC 37

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

ATGCGGCTGC ATAGCAACCT GACCAAAAAG 30

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

ATCCGCAGGT ACACCACAAC CATGATGAGG 30

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

TTTTGAGCAA GTTCAGCCTG GTTAAGT 27

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

TGGCTTATGA GTATTTCTTC CAGGGTA 27

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

GGGTAGTCGG TACCTCTAGA GCAAGTTCAG CC 32

```
(2) INFORMATION FOR SEQ ID NO:11:

    (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 31 base pairs
          (B) TYPE: nucleic acid
          (C) STRANDEDNESS: single
          (D) TOPOLOGY: linear

   (ii) MOLECULE TYPE: DNA

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

ATAACAGAGG ATCCTCGAGT ATTTCTTCCA G                                   31


(2) INFORMATION FOR SEQ ID NO:12:

    (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 1523 base pairs
          (B) TYPE: nucleic acid
          (C) STRANDEDNESS: single
          (D) TOPOLOGY: linear

   (ii) MOLECULE TYPE: DNA

   (ix) FEATURE:
          (A) NAME/KEY: CDS
          (B) LOCATION: 261..1322

   (ix) FEATURE:
          (A) NAME/KEY: Termination codon
          (B) LOCATION: 1320..1322

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

CACCTTCCTA ACCTGAGCGG CCTAGCCTGG GAAACAAACA ATTAAAATGT GCGCTAAATG     60

CTGTGGTAGG AGGTCAGGGG CTATGTCCTG GACCAAAGGA CATTTGCACT GAGACCTGAC    120

ACTTCAGGTC TTCAACTCCC TTGATGGGAG TTAGCCAGAA CGGGCTTAGA AACAGCAATT    180

GATGGCTTAG TGACTGATTT TACAAATGAT ATTTGTTTCT TCTTTAAATT TCTTTCTAGG    240

ATGTTCACTT CTTCTCCACA ATG AAT GAG TGT CAC TAT GAC AAG CAC ATG        290
                      Met Asn Glu Cys His Tyr Asp Lys His Met
                                  215                 220

GAC TTT TTT TAT AAT AGG AGC AAC ACT GAT ACT GTC GAT GAC TGG ACA      338
Asp Phe Phe Tyr Asn Arg Ser Asn Thr Asp Thr Val Asp Asp Trp Thr
            225                 230                 235

GGA ACA AAG CTT GTG ATT GTT TTG TGT GTT GGG ACG TTT TTC TGC CTG      386
Gly Thr Lys Leu Val Ile Val Leu Cys Val Gly Thr Phe Phe Cys Leu
    240                 245                 250

TTT ATT TTT TTT TCT AAT TCT CTG GTC ATC GCG GCA GTG ATC AAA AAC      434
```



```
Phe Ile Phe Phe Ser Asn Ser Leu Val Ile Ala Ala Val Ile Lys Asn
    255                 260                 265

AGA AAA TTT CAT TTC CCC TTT TAC TAC CTG TTG GCT AAT TTA GCT GCT      482
Arg Lys Phe His Phe Pro Phe Tyr Tyr Leu Leu Ala Asn Leu Ala Ala
270                 275                 280                 285

GCC GAT TTC TTC GCT GGA ATT GCC TAT GTA TTC CTG ATG TTT AAC ACA      530
Ala Asp Phe Phe Ala Gly Ile Ala Tyr Val Phe Leu Met Phe Asn Thr
                290                 295                 300

GGC CCA GTT TCA AAA ACT TTG ACT GTC AAC CGC TGG TTT CTC CGT CAG      578
Gly Pro Val Ser Lys Thr Leu Thr Val Asn Arg Trp Phe Leu Arg Gln
            305                 310                 315

GGG CTT CTG GAC AGT AGC TTG ACT GCT TCC CTC ACC AAC TTG CTG GTT      626
Gly Leu Leu Asp Ser Ser Leu Thr Ala Ser Leu Thr Asn Leu Leu Val
        320                 325                 330

ATC GCC GTG GAG AGG CAC ATG TCA ATC ATG AGG ATG CGG GTC CAT AGC      674
Ile Ala Val Glu Arg His Met Ser Ile Met Arg Met Arg Val His Ser
    335                 340                 345

AAC CTG ACC AAA AAG AGG GTG ACA CTG CTC ATT TTG CTT GTC TGG GCC      722
Asn Leu Thr Lys Lys Arg Val Thr Leu Leu Ile Leu Leu Val Trp Ala
350                 355                 360                 365

ATC GCC ATT TTT ATG GGG GCG GTC CCC ACA CTG GGC TGG AAT TGC CTC      770
Ile Ala Ile Phe Met Gly Ala Val Pro Thr Leu Gly Trp Asn Cys Leu
                370                 375                 380

TGC AAC ATC TCT GCC TGC TCT TCC CTG GCC CCC ATT TAC AGC AGG AGT      818
Cys Asn Ile Ser Ala Cys Ser Ser Leu Ala Pro Ile Tyr Ser Arg Ser
            385                 390                 395

TAC CTT GTT TTC TGG ACA GTG TCC AAC CTC ATG GCC TTC CTC ATC ATG      866
Tyr Leu Val Phe Trp Thr Val Ser Asn Leu Met Ala Phe Leu Ile Met
        400                 405                 410

GTT GTG GTG TAC CTG CGG ATC TAC GTG TAC GTC AAG AGG AAA ACC AAC      914
Val Val Val Tyr Leu Arg Ile Tyr Val Tyr Val Lys Arg Lys Thr Asn
    415                 420                 425

GTC TTG TCT CCG CAT ACA AGT GGG TCC ATC AGC CGC CGG AGG ACA CCC      962
Val Leu Ser Pro His Thr Ser Gly Ser Ile Ser Arg Arg Arg Thr Pro
430                 435                 440                 445

ATG AAG CTA ATG AAG ACG GTG ATG ACT GTC TTA GGG GCG TTT GTG GTA     1010
Met Lys Leu Met Lys Thr Val Met Thr Val Leu Gly Ala Phe Val Val
                450                 455                 460

TGC TGG ACC CCG GGC CTG GTG GTT CTG CCC CTC GAC GGC CTG AAC TGC     1058
Cys Trp Thr Pro Gly Leu Val Val Leu Pro Leu Asp Gly Leu Asn Cys
            465                 470                 475

AGG CAG TGT GGC GTG CAG CAT GTG AAA AGG TGG TTC CTG CTG CTG GCG     1106
Arg Gln Cys Gly Val Gln His Val Lys Arg Trp Phe Leu Leu Leu Ala
        480                 485                 490

CTG CTC AAC TCC GTC GTG AAC CCC ATC ATC TAC TCC TAC AAG GAC GAG     1154
Leu Leu Asn Ser Val Val Asn Pro Ile Ile Tyr Ser Tyr Lys Asp Glu
```

```
495                 500                 505

GAC ATG TAT GGC ACC ATG AAG AAG ATG ATC TGC TGC TTC TCT CAG GAG      1202
Asp Met Tyr Gly Thr Met Lys Lys Met Ile Cys Cys Phe Ser Gln Glu
510                 515                 520                 525

AAC CCA GAG AGG CGT CCC TCT CGC ATC CCC TCC ACA GTC CTC AGC AGG      1250
Asn Pro Glu Arg Arg Pro Ser Arg Ile Pro Ser Thr Val Leu Ser Arg
                530                 535                 540

AGT GAC ACA GGC AGC CAG TAC ATA GAG GAT AGT ATT AGC CAA GGT GCA      1298
Ser Asp Thr Gly Ser Gln Tyr Ile Glu Asp Ser Ile Ser Gln Gly Ala
            545                 550                 555

GTC TGC AAT AAA AGC ACT TCC TAA ACTCTGGATG CCTCTYGGCC CACCCAGGCC     1352
Val Cys Asn Lys Ser Thr Ser
    560                 565

TCCTCTGGGA AAAGAGCTGT TAAGAATGAT TACCTGTCTC TAACAAAGCC CATGTACAGT    1412

GTTATTTGAG GTCTCCATTA ATCACTGCTA GATTTCTTTA AAAAATTTTT TTTCATAGTT    1472

TAAAAGCATG GGCAGTAAAG AGAGGACCTG CTGCATTTAG AGAAAGCACA G             1523
```

(2) INFORMATION FOR SEQ ID NO:13

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1356 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13

```
GAATTCGCGG CCGCGTCGAC GTTCACTTCT CCACAATGAA TGAGTGTCAC TATGACAAGC      60

ACATGGACTT TTTTTATAAT AGGAGCAACA CTGATACTGT CGATGACTGG ACAGGAACAA     120

AGCTTGTGAT TGTTTTGTGT GTTGGGACGT TTTTCTGCCT GTTTATTTTT TTTTCTAATT     180

CTCTGGTCAT CGCGGCAGTG ATCAAAAACA GAAAATTTCA TTTCCCCTTT TACTACCTGT     240

TGGCTAATTT AGCTGCTGCC GATTTCTTCG CTGGAATTGC CTATGTATTC CTGATGTTTA     300

ACACAGGCCC AGTTTCAAAA ACTTTGACTG TCAACCGCTG GTTTCTCCGT CAGGGGCTTC     360

TGGACAGTAG CTTGACTGCT TCCCTCACCA ACTTGCTGGT TATCGCCGTG GAGAGGCACA     420

TGTCAATCAT GAGGATGCGG GTCCATAGCA ACCTGACCAA AAAGAGGGTG ACACTGCTCA     480

TTTTGCTTGT CTGGGCCATC GCCATTTTTA TGGGGGCGGT CCCCACACTG GGCTGGAATT     540

GCCTCTGCAA CATCTCTGCC TGCTCTTCCC TGGCCCCCAT TTACAGCAGG AGTTACCTTG     600

TTTTCTGGAC AGTGTCCAAC CTCATGGCCT TCCTCATCAT GGTTGTGGTG TACCTGCGGA     660

TCTACGTGTA CGTCAAGAGG AAAACCAACG TCTTGTCTCC GCATACAAGT GGGTCCATCA     720
```



```
GCCGCCGGAG GACACCCATG AAGCTAATGA AGACGGTGAT GACTGTCTTA GGGGCGTTTG     780

TGGTATGCTG GACCCCGGGC CTGGTGGTTC TGCCCCTCGA CGGCCTGAAC TGCAGGCAGT     840

GTGGCGTGCA GCATGTGAAA AGGTGGTTCC TGCTGCTGGC GCTGCTCAAC TCCGTCGTGA     900

ACCCCATCAT CTACTCCTAC AAGGACGAGG ACATGTATGG CACCATGAAG AAGATGATCT     960

GCTGCTTCTC TCAGGAGAAC CCAGAGAGGC GTCCCTCTCG CATCCCCTCC ACAGTCCTCA    1020

GCAGGAGTGA CACAGGCAGC CAGTACATAG AGGATAGTAT TAGCCAAGGT GCAGTCTGCA    1080

ATAAAAGCAC TTCCTAAACT CTGGATGCCT CTGGCCCACC CAGGCCTCCT CTGGGAAAAG    1140

AGCTGTTAAG AATGATTACC TGTCTCTAAC AAAGCCCATG TACAGTGTTA TTTGAGGTCT    1200

CCATTAATCA CTGCTAGATT TCTTTAAAAA ATTTTTTTTC ATAGTTTAAA AGCATGGGCA    1260

GTAAAGAGAG GACCTGCTGC ATTTAGAGAA AGCACAGGTC GACGCGGCCG CGAATTCTTT    1320

TGCTTTTTAC CCTGGAAGAA ATACTCGAGC ATGCAT                              1356
```

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 353 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

```
Met Asn Glu Cys His Tyr Asp Lys His Met Asp Phe Phe Tyr Asn Arg
1               5                   10                  15

Ser Asn Thr Asp Thr Val Asp Asp Trp Thr Gly Thr Lys Leu Val Ile
            20                  25                  30

Val Leu Cys Val Gly Thr Phe Phe Cys Leu Phe Ile Phe Phe Ser Asn
        35                  40                  45

Ser Leu Val Ile Ala Ala Val Ile Lys Asn Arg Lys Phe His Phe Pro
    50                  55                  60

Phe Tyr Tyr Leu Leu Ala Asn Leu Ala Ala Ala Asp Phe Phe Ala Gly
65                  70                  75                  80

Ile Ala Tyr Val Phe Leu Met Phe Asn Thr Gly Pro Val Ser Lys Thr
                85                  90                  95

Leu Thr Val Asn Arg Trp Phe Leu Arg Gln Gly Leu Leu Asp Ser Ser
            100                 105                 110

Leu Thr Ala Ser Leu Thr Asn Leu Leu Val Ile Ala Val Glu Arg His
        115                 120                 125

Met Ser Ile Met Arg Met Arg Val His Ser Asn Leu Thr Lys Lys Arg
```

```
        130                 135                 140

Val Thr Leu Leu Ile Leu Leu Val Trp Ala Ile Ala Ile Phe Met Gly
145                 150                 155                 160

Ala Val Pro Thr Leu Gly Trp Asn Cys Leu Cys Asn Ile Ser Ala Cys
                165                 170                 175

Ser Ser Leu Ala Pro Ile Tyr Ser Arg Ser Tyr Leu Val Phe Trp Thr
            180                 185                 190

Val Ser Asn Leu Met Ala Phe Leu Ile Met Val Val Val Tyr Leu Arg
        195                 200                 205

Ile Tyr Val Tyr Val Lys Arg Lys Thr Asn Val Leu Ser Pro His Thr
    210                 215                 220

Ser Gly Ser Ile Ser Arg Arg Arg Thr Pro Met Lys Leu Met Lys Thr
225                 230                 235                 240

Val Met Thr Val Leu Gly Ala Phe Val Val Cys Trp Thr Pro Gly Leu
                245                 250                 255

Val Val Leu Pro Leu Asp Gly Leu Asn Cys Arg Gln Cys Gly Val Gln
            260                 265                 270

His Val Lys Arg Trp Phe Leu Leu Leu Ala Leu Leu Asn Ser Val Val
        275                 280                 285

Asn Pro Ile Ile Tyr Ser Tyr Lys Asp Glu Asp Met Tyr Gly Thr Met
    290                 295                 300

Lys Lys Met Ile Cys Cys Phe Ser Gln Glu Asn Pro Glu Arg Arg Pro
305                 310                 315                 320

Ser Arg Ile Pro Ser Thr Val Leu Ser Arg Ser Asp Thr Gly Ser Gln
                325                 330                 335

Tyr Ile Glu Asp Ser Ile Ser Gln Gly Ala Val Cys Asn Lys Ser Thr
            340                 345                 350

Ser
```

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 213 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

```
Asn Thr Gly Pro Val Ser Lys Thr Leu Thr Val Asn Arg Trp Phe Leu
  1               5                  10                  15

Arg Gln Gly Leu Leu Asp Thr Ser Leu Thr Ala Ser Leu Ala Asn Leu
```



```
            20                  25                  30

Leu Val Ile Ala Val Glu Arg His Met Ser Ile Met Arg Met Arg Val
        35                  40                  45

His Ser Asn Leu Thr Lys Lys Arg Val Thr Leu Leu Ile Leu Leu Val
    50                  55                  60

Trp Ala Ile Ala Ile Phe Met Gly Ala Val Pro Thr Leu Gly Trp Asn
65                  70                  75                  80

Cys Leu Cys Asn Ile Ser Ala Cys Ser Ser Leu Ala Pro Ile Tyr Ser
                85                  90                  95

Arg Ser Tyr Leu Ile Phe Trp Thr Val Ser Asn Leu Leu Ala Phe Phe
            100                 105                 110

Ile Met Val Ala Val Tyr Val Arg Ile Tyr Met Tyr Val Lys Arg Lys
        115                 120                 125

Thr Asn Val Leu Ser Pro His Thr Ser Gly Ser Ile Ser Arg Arg Arg
    130                 135                 140

Ala Pro Met Lys Leu Met Lys Thr Val Met Thr Val Leu Gly Ala Phe
145                 150                 155                 160

Val Val Cys Trp Thr Pro Gly Leu Val Val Leu Leu Leu Asp Gly Leu
                165                 170                 175

Asn Cys Lys Gln Cys Asn Val Gln His Val Lys Xaa Trp Phe Leu Leu
            180                 185                 190

Leu Ala Leu Leu Asn Ser Val Met Asn Pro Leu Ile Tyr Cys Arg Ser
        195                 200                 205

Pro Xaa Phe Pro Trp
    210
```

INTERNATIONAL SEARCH REPORT

International Application No
PCT/CA 98/01193

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/12 C07K14/705 C12Q1/68 G01N33/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED
Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N C07K C12Q G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	EP 0 878 479 A (SMITHKLINE BEECHAM CORP/PLC (US); ELLIS; SATHE; VAN HORN; CLINKENBEARD) 18 November 1998 see the whole document ---	1-20
A	AN S. ET AL.: "Molecular cloning of the human EDG2 protein and its identification as a functional cellular receptor for lysophosphatidic acid" BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS, vol. 231, no. 3, 24 February 1997, pages 619-622, XP002046899 cited in the application see abstract --- -/--	5,19

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier document but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
7 May 1999	21/05/1999

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Macchia, G

1

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International Application No
PCT/CA 98/01193

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	AN S. ET AL.: "Identification of cDNAs encoding two G protein-coupled receptors for lysosphingolipids" FEBS LETTERS, vol. 417, no. 3, 17 December 1997, pages 279-282, XP002102182 see abstract ---	5,19
A	BURNSTOCK G.: "P2 purinoreceptors: historical perspective and classification" CIBA FOUNDATION SYMPOSIUM, vol. 198, 1996, pages 1-34, XP002102150 ---	
A	LEE N.H. ET AL.: "Molecular biology of G-protein-coupled receptors" DRUG NEWS AND PERSPECTIVES, vol. 6, no. 7, 1 September 1993, pages 488-497, XP000677175 ---	
A	OLIVEIRA L. ET AL.: "A common motif in G-protein-coupled seven transmembrane helix receptors" JOURNAL OF COMPUTER-AIDED MOLECULAR DESIGN, vol. 7, no. 6, 1 December 1993, pages 649-658, XP002050853 -----	

1

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/CA 98/01193

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 0878479 A	18-11-1998	CA 2230971 A	13-11-1998
		JP 11018788 A	26-01-1999